**EMOTION DETECTION SENTIMENT ANALYSIS: TEXT PROCESSING SENTIMENT ANALYSIS**

By

Sandra Tang Poh Yi



FACULTY OF COMPUTING AND

INFORMATION TECHNOLOGY

TUNKU ABDUL RAHMAN UNIVERSITY OF MANAGEMENT AND TECHNOLOGY

KUALA LUMPUR

ACADEMIC YEAR

**2024/25**

Emotion Detection Sentiment Analysis: Text Processing Sentiment Analysis

By

Sandra Tang Poh Yi

Supervisor: Anurehka A/P Magheswaran

A project report submitted to the

Faculty of Computing and Information Technology

in partial fulfillment of the requirement for the

Bachelor of Software Engineering (Honours)

**Department of Software Engineering and Technology**

Faculty of Computing and Information Technology

Tunku Abdul Rahman University of Management and Technology

Kuala Lumpur

Copyright by Tunku Abdul Rahman University of Management and Technology.

All rights reserved. No part of this project documentation may be reproduced, stored in retrieval system, or transmitted in any form or by any means without prior permission of Tunku Abdul Rahman University of Management and Technology.

# 

Declaration

The project submitted herewith is a result of my own efforts in totality and in every aspect of the project works. All information that has been obtained from other sources has been fully acknowledged. I understand that any plagiarism, cheating or collusion or any sorts constitutes a breach of TAR University rules and regulations and would be subjected to disciplinary actions.

A signature on a white background

Description automatically generated

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Sandra Tang Poh Yi

Bachelor of Software Engineering (Honours)

ID: 22WMR13625

Abstract

In the cyber era, however, social media has become a significant way of expressing emotions, but it is difficult to use them in mental health monitoring due to too much information. This project outlines a machine learning and natural language processing-based sentiment analysis platform that detects and analyzes text from social media, identifying feelings like distress, depression, and suicidal thoughts. The system rapidly notifies mental health practitioners thereby enabling them to interact personally with the respondents on web support through live chats thus providing immediate help based on individual needs. To reach out to those who may not seek traditional aid, the system is interfaced with mental healthcare services intervening promptly in such cases as this enhances early treatment and improves the mental health care system. On top of that, the platform embraces different linguistic and cultural aspects by supporting several languages whilst it keeps coping with evolving trends towards reducing mental health emergencies while enhancing wellness.

**Table of Contents**

[Declaration iii](#_Toc195891922)

[Abstract iv](#_Toc195891923)

[Table of Contents v](#_Toc195891924)

[1 Introduction 8](#_Toc195891925)

[1.1 Objective 8](#_Toc195891926)

[1.2 Project Background 8](#_Toc195891927)

[1.3 Advantages and Contribution 11](#_Toc195891928)

[1.3.1 Advantages 11](#_Toc195891929)

[1.3.2 Contribution 11](#_Toc195891930)

[1.4 Project Plan 13](#_Toc195891931)

[1.4.1 Project 1 14](#_Toc195891932)

[1.4.2 Project 2 14](#_Toc195891933)

[1.5 Project Team & Organization 15](#_Toc195891934)

[1.6 Chapter Summary and Evaluation 16](#_Toc195891935)

[2 Literature Review 18](#_Toc195891936)

[2.1 Background and Related Work 18](#_Toc195891937)

[2.2 Literature Review 22](#_Toc195891938)

[2.2.1 Lexicon-Based Approaches 22](#_Toc195891939)

[2.2.2 Machine Learning Approaches 23](#_Toc195891940)

[2.2.3 Deep Learning Approaches 23](#_Toc195891941)

[2.2.4 Cloud Infrastructure for Sentiment Analysis 24](#_Toc195891942)

[2.2.5 Comparison of Approaches 25](#_Toc195891943)

[2.2.6 Summarization Comparison Table 29](#_Toc195891944)

[2.2.7 Comparison Table of Algorithms 31](#_Toc195891945)

[2.2.8 Comparison of Tools 32](#_Toc195891946)

[2.2.9 Comparison of Cloud Platforms 33](#_Toc195891947)

[2.3 Feasibility Study 35](#_Toc195891948)

[2.3.1 Technical Feasibility 35](#_Toc195891949)

[2.3.2 Operational Feasibility 35](#_Toc195891950)

[2.3.3 Economic Feasibility 35](#_Toc195891951)

[2.3.4 Schedule Feasibility 36](#_Toc195891952)

[2.3.5 Legal Feasibility 36](#_Toc195891953)

[2.4 Chapter Summary and Evaluation 37](#_Toc195891954)

[3 Methodology and Requirements Analysis 40](#_Toc195891955)

[3.1 Overall Methodology 40](#_Toc195891956)

[3.2 Proposed Software Process Model 43](#_Toc195891957)

[3.3 System Requirements 45](#_Toc195891958)

[3.3.1 Functional Requirements 45](#_Toc195891959)

[3.3.2 Non-Functional Requirements 46](#_Toc195891960)

[3.4 Other Requirements 46](#_Toc195891961)

[3.4.1 Hardware Requirements 46](#_Toc195891962)

[3.4.2 Software Requirements 47](#_Toc195891963)

[3.5 Chapter Summary and Evaluation 48](#_Toc195891964)

[4 System Design 51](#_Toc195891965)

[4.1 Algorithm/Process Design 52](#_Toc195891966)

[4.1.1 Activity Diagram: Text Input and Preprocessing 52](#_Toc195891967)

[4.1.2 Activity Diagram: Sentiment and Emotion Detection 53](#_Toc195891968)

[4.2 Database Design/Structural Model 54](#_Toc195891969)

[4.2.1 ERD 54](#_Toc195891970)

[4.2.2 Class Diagram 55](#_Toc195891971)

[4.3 Interaction Model 57](#_Toc195891972)

[4.3.1 Text Sentiment Analysis Subsystem Main Use Case Diagram 57](#_Toc195891973)

[4.3.2 Use Case 1 – User Management Module 58](#_Toc195891974)

[4.3.3 Use Case 2 – Text Sentiment Analysis Module 61](#_Toc195891975)

[4.4 Behavioral Model 63](#_Toc195891976)

[4.4.1 User Management Module 63](#_Toc195891977)

[4.4.2 Text Sentiment Analysis Module 64](#_Toc195891978)

[4.5 Others Design - UI Design 65](#_Toc195891979)

[4.6 Chapter Summary and Evaluation 74](#_Toc195891980)

[5 Implementation and Testing 76](#_Toc195891981)

[5.1 Implementation 76](#_Toc195891982)

[5.1.1 Text Preprocessing 77](#_Toc195891983)

[5.1.2 Bidirectional LSTM Model Architecture 78](#_Toc195891984)

[5.1.3 Emotion Analyze Class 79](#_Toc195891985)

[5.1.4 Web Application Implementation 80](#_Toc195891986)

[5.2 Testing (Test Plan and Test Cases) 83](#_Toc195891987)

[5.2.1 Test Plan 83](#_Toc195891988)

[Overview 83](#_Toc195891989)

[Context of Testing 83](#_Toc195891990)

[Project 83](#_Toc195891991)

[Test Item 83](#_Toc195891992)

[Test Scope 84](#_Toc195891993)

[Assumption and Constraints 86](#_Toc195891994)

[Risk Register 86](#_Toc195891995)

[Test Strategy 87](#_Toc195891996)

[Test Completion Criteria 88](#_Toc195891997)

[Metrics to be collected 89](#_Toc195891998)

[Test Data Requirement 89](#_Toc195891999)

[Test Environment Requirement 92](#_Toc195892000)

[Testing activities and estimates 93](#_Toc195892001)

[5.2.2 Test Cases 94](#_Toc195892002)

[5.3 Chapter Summary and Evaluation 111](#_Toc195892003)

[6 Discussions and Conclusion 113](#_Toc195892004)

[6.1 Summary 113](#_Toc195892005)

[6.2 Achievements 113](#_Toc195892006)

[6.3 Contributions 114](#_Toc195892007)

[6.4 Limitations and Future Improvements 115](#_Toc195892008)

[6.5 Issue and Solutions 115](#_Toc195892009)

[References 117](#_Toc195892010)

Chapter 1

Introduction

# Introduction

In this chapter, the project as a whole is discussed, including its goals, context, benefits and contributions. It describes the project management including activities, schedules, and deadlines, the organization of the project and assignment of the modules. At the end of the chapter, there is a brief conclusion and assessment in order to emphasize the main points of the chapter.

## Objective

1. To investigate the emotions that are communicated through text.
2. To gather textual information, process it, identify and evaluate emotions.
3. To develop a text-based sentiment analysis and emotion detection system.

## Project Background

With the rapid growth of social networking sites in this technological era, these platforms have become indispensable for sharing emotions with people worldwide. A lot of individuals convey their opinions and emotions through text, visuals, videos and audio recording in social media (Nandwani & Verma, 2021). Thus, it is possible to research how people behave using digital media (Tiwari et al., 2021). Other than that, it’s been reported recently on online platforms that an increasing number of young people are daily dealing with mental health issues especially stress and depression (Shayaa et al., 2018). For example, anxiety and depressive disorders among students are primary caused by their studies particularly during the COVID-19 pandemic when educational institutions are close down and virtual learning has to be put into place (Syafirah et al., 2022). In addition, academic issues are one of the things that contribute to student’s anxiety and sadness. For instance, anxiety about failing, decline in assessments and intense competition (Syafirah et al., 2022).

Besides that, amidst this worldwide heath crisis, residents are subjected to lockdown and quarantine, that hinders their daily routines. When there are no social contact and people witness the death of someone, they care about are more likely to experience feelings of anxiety and depression which can lead to persistent mental disorders (Banna et al., 2023). As a result, more people are using social media to discuss their mental health issues and share their opinions about the COVID-19 pandemic. In addition, people who suffer from depression prefer to share their emotions on social networking sites because they shy away from social situations (Banna et al., 2023). According to research, people with a variety of mental conditions, such as depressive disorders, psychotic disorders, or other serious mental disorders, use social networking sites at rates similar to the entire population. Among middle-aged and older people, use of social media ranges from approximately 70 percent to 97 percent (Aschbrenner et al., 2018, Naslund et al., 2020). Moreover, Wang talks about the dual nature of social media, pointing out that it can help people make online friends but can also cause feelings of isolation, depression, and negative body image because of peer comparisons (Rizal et al., n.d.). Other than that, COVID-19 has caused a significant rise in Post-traumatic Stress Disorder (PTSD) (Sawalha et al., 2022). According to a survey of 24 countries, 70% of people will encounter at least one potentially traumatic event (PTE) during their entire lives prior to the worldwide outbreak (Benjet et al., 2016). We must enhance detection and evaluation methods in light of the increasing incidence of PTSD, particularly in light of the recent development of telepsychiatry (Sawalha et al., 2022).

The project will develop a comprehensive sentiment analysis platform designed to analyse text data from social media, online reviews and other digital content. This platform will identify and interpret emotions expressed in textual content by utilizing cutting-edge machine learning and natural language processing (NLP) techniques. As the platform monitors social media and online platforms in real-time, it will alert users if there is content that indicates distress, depression, or suicidal thoughts. Thus, this service allows the mental health professionals and relevant authorities to respond promptly and offer necessary interventions. Other than that, it can have a multilingual support which can support multiple language including Malay, English, Chinese. It will be built to comprehend context-specific language and cultural quirks in order to reliably translate sentiment across a range of linguistic environments. Moreover, by identifying people who are at risk and offering insights into public sentiment trends, the platform is intended to work in tandem with mental health services. In addition, with this collaboration, mental health support and outreach are aimed at those who do not seek traditional mental health treatment. However, the services provided by the solution is emotional sentiment analysis which can analyse text data to detect emotions such as happiness, sadness, anger and fear. Other than that, it will have a real-time alert which can send out immediate alerts for any content that might be harmful to someone’s mental health. Besides that, it will have a language and cultural adaption which can support multiple languages and understand cultural nuances in emotional expression. Furthermore, it also includes customer insight report that can let customer to generate reports on customer sentiment trends and feedback from therapy. Lastly, it can also share insights with mental health services to facilitate early intervention and support.

Other than that, real-time monitoring and alerts from the platform allow mental health professionals to intervene promptly, potentially saving lives and offering essential support. In addition to supporting multiple languages, the platform offers meaningful insights by capturing sentiment across Malaysia's diverse population. Moreover, the platform has the ability to identify users who are expressing distress online and direct them to relevant mental health resources and support networks. Directing them to helplines, counselling services, and online support communities may be one way to reach out to those who have not sought assistance before. Besides that, the platform will use machine learning techniques to make constant improvements to its algorithms, making sure that it remains current with changing mental health trends and language patterns. Hence, this flexibility makes the platform capable of identifying mental health problems even in the face of shifting social norms and communication preferences.

## Advantages and Contribution

### Advantages

Emotion detection through text sentiment analysis is very beneficial to many people especially those battling with various mental health problems. Text sentiment analysis allows identifying emotions in text documents from social networks, forums and other sites such as signs of stress or depression, suicidal tendencies, etc. in real-time. This helps therapists and organizations to reach out to people who are at risk yet most likely would not seek therapy in the traditional sense. Thus, the early recognition of depression is very important for intervention and potential saving of lives and limiting the spending of health care systems. Moreover, through the use of sentiment analysis, large groups of individuals can be targeted assisting with mental health issues, particularly in areas with few mental health services. And also, although it is not only confined to mental health, sentiment and emotion analysis aids in softening various relationships in order to avoid negative impacts in situations like consumer care, teaching or even governance by detecting emotions that are not usually verbalized. This technology in the end works towards a more humanistic and kind society that benefits the health of all people.

### Contribution

The creation of software that makes use of text-based emotion detection has the potential to greatly benefit a number of industries, including organizations, mental health services, social media and private citizens. For organizations, it can implement emotion detection tools to monitor and improve employee mental health and well-being. A timely intervention can reduce burnout and improve productivity when stress or dissatisfaction are detected early. Other than that, companies can better deliver services by understanding emotions through the analysis of customer interactions. By using this software, it can effectively manage customer emotions which can help businesses improve customer satisfaction and loyalty that can lead to boost sales and retention. Besides that, when text-based emotion detection software is used in conjunction with mental health services, its benefits can be greatly increased, which is advantageous to both the service providers and the people they assist. Professionals in mental health gain a valuable tool to improve their diagnostic abilities by incorporating emotion detection technology. In order to identify emotional patterns and possible mental health concerns, the software can examine text inputs from patients, such as messages or replies in therapeutic chatbots. Thus, this aids medical professionals in recognizing symptoms that conventional approaches might miss. As a result, this will lead to more accurate diagnoses and personalized treatment plans and improve the overall effectiveness of therapy and care. On the other hand, an emotion detection software can determine a person's emotional state and sentiments based on the text in their social media posts, comments, and messages. This analysis can identify mood swings, indicators of mental health problems, or distress signals from both public and private messages. Hence, this enables early detection of individuals who may need mental health support. Other than that, social media can also collaborate with mental health services. Social media data may give mental health providers important information about people's emotional health as well as that of their communities. Through this partnership, mental health services can now provide more coordinated care by customizing their outreach and interventions according to real-time social media analysis. As a result, it improves mental health services efficacy by incorporating outside data sources. Other than that, people can learn more about their emotional health by using emotion detection technologies. This is because it can increase self-awareness and access to coping strategies that can lead to improve mental health. Moreover, personalized suggestions for self-care resources and activities can be obtained through emotion detection. As a result, individuals receive advice tailored to their current emotional state, which helps them choose healthier lifestyles.

## Project Plan

|  |  |  |
| --- | --- | --- |
| **ACTIVITIES** | **EXPECTED OUTCOME** | **COMPLETION DATE** |
| Proposal Writing | Detailed project plan, including scope, timeline, resources, and deliverables | 18/8/2024 |
| Introduction | Project Introduction (Chapter 1) | 24/9/2024 |
| Research Background | Research Background Documentation (Chapter 2) | 17/11/2024 |
| Analyze method and requirements | Methodology and Requirements Analysis (Chapter 3) | 10/12/2024 |
| Design the System | System Design (Chapter 4) | 17/12/2024 |
| System Preview | Source code | 24/02/2025 |
| System Testing | Final source code | 10/03/2025 |
| Draft FYP Report | Workable Text Sentiment Analysis System | 07/04/2025 |
| Final FYP Report | Completed Text Sentiment Analysis System | 21/04/2025 |

**Table 1.1 Project Plan Table**

The Project Plan Table (Table 1.1) describes the key dates and milestones necessary to successfully accomplish the project of text sentiment analysis project. It includes tasks such as preparing a proposal, conducting research background, formulating methodology and requirements analysis, system design and testing. Every phase of the project has its own sets of outputs and timeframes, starting from Proposal Writing on 18 August 2024 and closing with the Final FYP Report on 21 April 2025. This structured plan greatly facilitates the orderly progression and timely completion of the project.

**Figure 1.1 Project Plan Gantt Chart**

### A grid of lines and numbers Description automatically generated with medium confidenceProject 1

*A grid with a blue rectangle

Description automatically generatedA grid with a blue line

Description automatically generated*

### Project 2

*A grid of white squares

Description automatically generatedA grid of squares with a blue line

Description automatically generated with medium confidenceA grid of black lines

Description automatically generated*

The Project Plan Gantt Chart allows a clear understanding of the sequence of events that took place during the execution of the project ranging from the project proposal to the achievement and submission of the project report. It shows graphically when each of the activities within a project begins and ends, showing linkages and proper scheduling in order to accomplish the aims of the project.

This is the access link to the Gantt Chart: [Gantt Chart.xlsx](https://kutar-my.sharepoint.com/:x:/g/personal/2213626_m365_tarc_edu_my/EaHGCTnADoxEuEkKMZ-DDxYB1_hRBmTzQJmDpYbDq8stvA?e=89vspo)

## Project Team & Organization

|  |  |  |
| --- | --- | --- |
| Module | Sandra Tang Poh Yi | Saw Hui Lin |
| Data Collection  (procedure for collecting and evaluating data on the relevant variable) | x | x |
| Text Processing  (Text data is analysed and transformed, using sentiment analysis to find pertinent details, structures, or ideas.) | x |  |
| Audio Processing   (Prepare audio data for analysis) |  | x |
| Feature Extraction  (extract relevant features from the processed data) | x | x |
| System Testing  (ensure the system meets the required standards) | x | x |

**Table 1.2 Project Team and Modules**

This is our project team, and we have the list of modules here. So, I deal with tasks such as preparing, gathering, encoding, and evaluating the text and feature extraction as well as testing the system. Moreover, it’s also known that data in regard to audio is processed by Saw Hui Lin as well, she participates in such activities as gathering, encoding and evaluating the data and the system. As a group, we work together on certain key blocks which are essential for the timely creation and deployment of the system.

## Chapter Summary and Evaluation

The implementation of the text analytic program focused on sentiment analysis is a step forward in tackling the increasing challenges faced by people suffering with mental health. The platform utilizes machine learning, natural language processing (NLP), and other real time monitoring tools to help identifying detrimental states like depression, stress or suicidal ideation in online content. This innovation assists mental health professionals in delivering timely care and allows organizations, social media sites and individuals to learn about emotional health affairs. The platform offers cultural and linguistic relevance to cater to the needs of different population groups and thus encourages a more caring and informed community.

Chapter 2

Literature Review

# Literature Review

This chapter analyzes previous studies, similar techniques, and software relevant measuring feelings or emotions and text evaluation. It deals with the current problems, evaluates approaches and gives compromises in the course of making the study. Finally, this chapter presents a conclusion and assessment of the results.

## Background and Related Work

Individual personalities and behavioral characteristics are reflected in the complex, multifaceted quality of emotion. People express their feelings about a variety of topics, including events, people, the environment and even the smallest things that are in their immediate vicinity, in the course of their daily lives (Sailunaz & Alhajj, 2019). Emotion detection and sentiment analysis have evolved into a critical area of research in the realm of natural language processing (NLP). This field focuses on interpreting human emotions from textual data, enabling machines to analyse and understand emotions with a wide range of applications. Other than that, sentiment analysis has become an increasingly common method to get deeper understanding, like figuring out how happy our people are, imagining how consumers feel about the products or services which run by multinational companies, predicting outcome of elections and movements in the stock market from Twitter data and estimating box office receipts from movie reviews (Salam & Gupta, 2018).

Moreover, the capacity to gauge emotions from text has been particularly beneficial in areas such as customer service, education, mental healthcare, politics, and social media analysis, as it allows for real-time sentiment tracking and emotional insights (Nandwani & Verma, 2021). The exponential growth of user-generated data on platforms like social media has contributed significantly to the development of emotion detection systems, which leverage machine learning, deep learning, and rule-based models to interpret human emotions (Bharti et al., 2022). Furthermore, opinion analysis and early detection of depression are two more tasks that emotion detection can help with (Hung & Alias, 2023). Besides that, scholars have been working on identifying and evaluating emotions in text from a variety of disciplines including psychology, computer science, artificial intelligence and so on. Thus, to identify the appropriate emotion in text, a variety of techniques have been used (Sailunaz & Alhajj, 2019).

Apart from that, using text mining and natural language processing, sentiment analysis locates and removes subjective data from the written content (Wankhade et al., 2022). The intricacy of emotions in people makes it difficult to identify the right emotion in text due to there are many obstacles to be conquered. When a single text piece expresses several emotions, it becomes more challenging to identify those emotions from the text. Hence, emotion in writing can occasionally be so inherent that it is practically impossible to detect automatically. Many arrogant texts are frequently hard for other people to identify, much less for an algorithm to pick up on (Sailunaz & Alhajj, 2019). According to the scenario and field in which they are used, many words in various languages have different meanings and orientations (Wankhade et al., 2022). As a result, not many resources and tools are accessible for every language. Hence, two of the most important issues that scholars have recently focused on are sarcasm and irony (Wankhade et al., 2022). Sarcasm in language can be identified using a variety of techniques (Castro et al., 2019). However, comedy is highly culturally specific, and it is difficult for a machine to comprehend distinct cultural references, so the issue is far from being solved (Wankhade et al., 2022). The largest obstacle to all NLP tasks, including sentiment analysis, is informal writing style (Wankhade et al., 2022). When writing reviews or texts, people are very informal and frequently use difficult-to-understand acronyms, emojis, and shortcuts (Wankhade et al., 2022).

Besides that, a large number of user reviews and comment on different goods and services on different e-commerce websites. Customers' evaluations and ratings across various platforms motivate suppliers and service providers to improve their current products, services, or systems. Nowadays, practically every business or industry is going through a digital transformation, which has led to an increase in both structured and unstructured data (Nandwani & Verma, 2021). Thus, creating valuable insights from unstructured data to aid in decision-making is a massive task for businesses (Ahmad et al., 2020). For example, in the business sector, vendors effectively gather customer feedback and disseminate details regarding their products via social networking sites like Facebook, Instagram and so on (Israel Edem Agbehadji & Abosede Ijabadeniyi, 2021). Furthermore, Sentiment analysis helps marketers better understand the viewpoints of their customers so they can modify their goods and services as needed (Al Ajrawi et al., 2021). It is possible to observe how business and consumer sentiment affects the performance of stock markets in both developed and developing countries. The stock market may be somewhat impacted as a result, and investor sentiments influence their investment choices, which can quickly spread and intensify throughout the network (Ahmed, 2020). Sentiment and emotion analysis have therefore altered how we do business (Bhardwaj et al., 2015).

Healthcare professionals and citizens from across the country are increasingly using social media to provide information about health issues and connect with each other. Examples include Twitter, Facebook and other online social media sites (Garcia & Berton, 2021). However, patients' mental health suffered as a result of being told to avoid contact with their loved ones. Health professionals must employ automated sentiment and emotion analysis to protect patients from mental health conditions like depression (Singh et al., 2021). In order to prevent worsening mental health conditions, people frequently share their thoughts and feelings on websites through their posts. If someone appeared to be depressed, others could reach out to them for support. Moreover, both teachers and students benefit greatly from sentiment and emotion analysis in the educational field. The best way for educators to enhance their methods is to get immediate feedback from their pupils (Sangeetha & Prabha, 2020). In addition, the process of observing and drawing conclusions from open-ended textual feedback is difficult. The results of an emotion and sentiment analysis help organizations and educators make necessary corrections (Wankhade et al., 2022). Therefore, using sentiment and emotion analysis during the enrolment procedure can assist the student in choosing the best institution or instructor (Rao & Kishore Baglod, 2024).

On the other hand, the researchers have recently put forth a number of techniques to identify the emotions in a text, including learning-based, lexical affinity, keyword-based, and hybrid models (Acheampong et al., 2020). Initially, a rule-based method was introduced, which included two approaches: lexical affinity-based and keyword-based. Over time, a new method emerged, known as the learning-based approach, which proved to be more accurate and yielded better results. In this approach, various models are used to detect emotions. Many researchers have also explored combining different methods to create hybrid models in pursuit of higher accuracy. According to studies, deep learning models tend to outperform machine learning models when handling large volumes of text or data. However, for smaller datasets, machine learning models offer better accuracy. Despite these advancements, no approach has yet provided a fully reliable solution for detecting emotions in text (Bharti et al., 2022). Besides, based on the ISEAR dataset, Razek and Frasso have constructed a hierarchical tree to detect feelings in texts using dominant meaning method (Razek & Frasson, 2017). As a result of reviewing existing techniques for detecting emotions, Jain and Kulkarni have developed a vector space model (Jain & Kulkarni, 2014). Moreover, according to Thomas, he had employed the Multinomial Naïve Bayes classifier and Weighted Log-Likelihood Score (WLLS) to derive sentiment from sentences (Alswaidan & Menai, 2020).

Aside from that, according to the literature, the current methods for identifying emotions in online social networks can be separated in two main categories which are lexicon-based techniques and machine learning-based methods (Kramer, 2010). The lexicon-based methods utilized dictionaries like the linguistic inquiry and word count (LIWC) to extract emotional keywords. According to Coviello, the mental states of the posted posts were measured using LIWC and found psychological transmission in massive online social networks (Coviello et al., 2014). Using LIWC, Golder and Macy assessed both the beneficial and detrimental impacts of tweets and found intermittent and daily mood patterns at the individual level in various cultures (Golder & Macy, 2011). Other than that, the use of machine learning and natural language processing (NLP) methods to identify emotions suggestive of suicidal behavior is gaining traction (Desmet & Hoste, 2013). Furthermore, recent research in sentiment analysis has expanded to various domains and languages (Pontiki et al., 2016). In one study, tweet data was classified using unsupervised learning methods, with a sentiment lexicon that identified different emoticons and emoji-based ideograms to determine sentiments (Wolny, 2016). A limited set of standard emoticons was extracted from Twitter data to analyse their corresponding sentiments. With the aid of several classification techniques, including the Random Forest (RF), SVM, Decision Tree, and Naive Bayes (NB) algorithms, a combination sentiment classification approach was used (Wan & Gao, 2015, (“An Algorithm and Method for Sentiment Analysis Using the Text and Emoticon,” 2020).

As we have seen, emotion detection and sentiment analysis have emerged as core subjects of study within NLP that have cross industries applicability including customer service, healthcare, education, social platforms, and business. With the ability to analyze text to extract emotions, machines can provide emotional feedback in real time to create tailored experiences and assist in various areas such as mental health care, customer review and education feedback. Nonetheless, despite the many advances achieved, gaps remain particularly in such areas as sarcasm, irony, and more colloquial language that makes emotion detection more precise difficult. Attempts have been made to boost accuracy using various models including rule-based, machine learning based and hybrid models. For large data sets, deep learning techniques outperform other approaches, however, for low datasets, the machine learning models provide better results. In any case, emotion detection is a still a work in progress, as no one individual technique has been successful in every case. The expansion of research in this area merging these different methods and suitable for detecting emotion in informal text and cultural factors can contribute to developing better emotion detection systems which expands their potential use in practice.

## Literature Review

Emotion detection in sentiment analysis has garnered substantial attention within the fields of natural language processing (NLP) and machine learning due to its wide range of applications, from social media monitoring to customer sentiment analysis (Batbaatar et al., 2019, Garg & Saxena, 2024). The overarching goal of emotion detection is to automatically identify and classify the emotional content of textual data, using a combination of methodologies, algorithms, and tools (Nandwani & Verma, 2021). In this review, we discuss the most common methodologies for text processing sentiment analysis, the algorithms typically employed, and the tools used for prototype development, along with their respective pros and cons.

### Lexicon-Based Approaches

Emotion detection in the context of sentiment analysis can use a number of techniques such as lexicon-based approach, machine learning and deep learning but these are applicable in different areas (Nandwani & Verma, 2021). One of the oldest methods which is also on the simplest is the lexicon-based method which uses a word list or dictionary containing word or emotive scores and assigns those to a target word (Britzolakis et al., 2020). VADER (Valence Aware Dictionary for Sentiment Reasoning) and SentiWordNet are among the most popular tools because they are easy to understand and easy to use (Hutto & Gilbert, 2014). For example, VADER is specifically designed for dealing with social media posts as it employs rules on negations, punctuation and capitalization which is informal language norms (Rani & Jan, 2020). Lexicon approaches are however found wanting on context comprehension such as in cases of sarcasm or idioms, and they also do not perform well with specialized vocabularies (Sanyal & Barai, 2021). In spite of these flaws, the ease of their application and their capabilities of processing in real time makes them suitable for situations where speed is taken to be more important than accuracy (Nguyen et al., 2019).

### Machine Learning Approaches

Machine learning-based approaches, on the other hand, focus on training models over labelled datasets so that they can spot patterns on data and extrapolate on new text. Commonly used algorithms are Naive Bayes, Support Vector Machines (SVM) and Random Forests(Aftab et al., 2017). A probabilistic classifier, Naive Bayes performs relatively well with text classification tasks when combined with feature extraction approaches such as TF-IDF (Term Frequency-Inverse Document Frequency) and has relatively low computational overhead (Nasim et al., 2017). Naive Bayes models’ reliance on the independence of features is often problematic, as most textual features are related (Mohaiminul & Sultana, 2018). SVM, however, is very proficient at distinguishing the required data by the utilization of a hyperplane and is one of the most competent text classifiers for high dimensional data, although its overall overhead is high for larger datasets (Singh et al., 2023). Random Forests, which is also an ensemble learning approach, builds multiple decision trees for higher accuracy but again, the computational cost of implementing these significantly rises with increasing dataset size (Aggarwal et al., 2021). Lexicon-based approaches may be superior in certain aspects, but the machine learning techniques are most often more flexible and robust, factoring the reliance on labelled data for training (Kabir et al., 2019).

### Deep Learning Approaches

Other than that, in the recent past, deep learning-based methods have become more widely used because of their capability to parse the structure and the context in the offerings. Recurrent neural networks (RNNs), Long short-term memory networks (LSTMs), Transformers and BERT (bidirectional encoder representations from transformers) have raised the bar in the noise language tasks performance (Aslam et al., 2022, Devlin et al., 2018, Zad et al., 2021). RNNs are effective in working on sequential data types but have skirts the problem of long-term dependency, which has been addressed by lattices through memory embedding (Li et al., 2018). This limitation has been addressed with memory mechanisms in architectures like LSTMs, which improve long-term dependency handling (Huang et al., 2022).

LSTM (Long Short-Term Memory) networks are extensively implemented in emotion detection because of their capability to capture long-term dependencies in the text. This is achieved using a complex gating structure—comprised of input, forget, and output gates—each of which manages a particular function within the cell's information. In context, LSTMs can capture emotionally relevant information while discarding much of the irrelevant detail.

Another developed version is Bidirectional LSTM (BiLSTM), which processes sequences in both forward and reverse directions improving its ability to capture context from preceding as well as succeeding words. Having access to both ends is advantageous and significantly improves BiLSTM's understanding of the emotional subtleties embedded in the text. There are many studies reporting the effectiveness of this model on emotion detection tasks. for instance, BiLSTM was used effectively to capture nuanced emotions from social media posts by Sailunaz and Alhajj (2019) who reported better performance than prevailing methods. In other works, Kim et al. (2018) added attention components to BiLSTM models focusing the model on emotionally significant words in a sentence which improved the performance.

With regard to LSTM-based models, their strongest asset is the ability to identify subtle and context-dependent emotions that may be expressed across long sequences of natural language. Complementary advantages, however, are offered by Convolutional Neural Networks (CNNs). Emotionally relevant features and patterns are often localized, and CNNs have demonstrated, through more efficient computation, that capturing such features through convolutional filters can be highly effective. Research comparing the two has shown that while contextually dependent long-range relations are better handled by LSTMs, short-term or localized emotional cues are best executed by CNNs, sometimes leading to dominating results (Zhou et al., 2020).

More advanced transformer-based models that include BERT are even able to capture the relation between the words in the context of the sentence by looking at it from both front and back views (Devlin et al., 2018). This makes BERT deal with more intricate emotions and complex features of languages including the use of idioms and sarcasm much more efficiently (Cui et al., 2020). Regardless of the fact that such models obtain high accuracy, deep learning models are hostile to resources and usually trained and inferred on special hardware such as gpus or tpus (Salehinejad et al., 2018). They are also more difficult to interpret than traditional machine learning models due to their complexity (Freitag et al., 2018).

### Cloud Infrastructure for Sentiment Analysis

Sentiment analysis systems are now much more scalable and efficient which need to thanks to recent developments in cloud computing platforms like Microsoft Azure, Google Cloud Platform (GCP), and Amazon Web Services (AWS). By providing strong computational resources, pre-built AI services, and deployment capabilities, these platforms tackle the resource-intensive nature of sophisticated models like transformers. Moreover, cloud computing supports deep learning models such as ConvLSTMConv, which utilize distributed resources for the management of extensive datasets and their accuracy (Ghorbani et al., 2020). The learner’s feedback in e-learning systems is examined with the help of cloud sentiment analysis, portraying the advantages of cloud environment (Zarra et al., 2016). Furthermore, frameworks like the Tweet Sentiment Analysis (TSA) employ cloud technologies to evaluate consumer opinions on cloud services (Karamitsos et al., 2019).

Services like the AWS Comprehend or Google Natural Language API provide ready-made solutions for implementing sentiment analysis, that can be done in real-time (Opara, 2022). There have been an increasing number of studies comparing different cloud providers such as Azure and AWS in showing how sentiment analysis can predict how good a provider’s reputation is with the help of machine learning techniques (Qaisi & Aljarah, 2016). Analysis of sentiment for Big Data made to monitor brands makes use of the cloud in order to ease the high volume of social media data into a small space (Benedetto & Tedeschi, 2016). Medical sentiment analysis is also cloud based using safe and hardware accelerated systems for effective patient feedback evaluation (Torres et al., 2018).

Saleh et al. (2022) highlight the benefits of hybrid models such as CBiLSTM, which utilize cloud computing to automatically evaluate the reputation of other cloud services through their sentiment. According to Tamrakar & Madhavi (2022), tools such as Google Cloud or BigQuery reduce the entry barriers on performing sentiment analysis for large programs. Arcila Calderón et al., 2019 argue that distributed machine learning frameworks based on cloud systems are convenient for the streaming big data and real time sentiment analysis Distributed machine learning frameworks enhance operational efficiency and resolve issues of scalability and cost-effectiveness which makes them suitable for applications that require a lot of resources (García et al, 2020).

### Comparison of Approaches

A general narrative of lexicon-based, machine, and deep learning approaches reveals that they can offer unique benefits and shortcomings depending on the specific problem at hand. For instance, Lexicon-based approaches including VADER can be particularly beneficial for use cases such as monitoring social media and where speed and efficiency are paramount (Hutto & Gilbert, 2014). However, they can underperform for large or domain-specific datasets as they cannot contextualize the data (Min & Zulkarnain, 2020). Cloud platforms can enhance lexicon-based approaches by providing scalable APIs that process large datasets more efficiently, such as Google Cloud Natural Language API and AWS Comprehend (Opara, 2022). Machine learning-based methods such as Naive Bayes and SVM are less bounded by resources compared to lexicon-based methods but are more resource-efficient relatively (Aftab et al., 2017). These techniques work well for labelled datasets and organized issues (Pranckevičius & Marcinkevičius, 2017). Used transformers, such as BERT, achieve the best accuracy and understanding of multilayer textual relations but increase costs due to the complexity of models and their interpretability (Devlin et al., 2018). Apart from that, cloud-native platforms, like AWS Sagemaker and Google AI Platform, address these challenges by offering distributed training environments, automated preprocessing, and real-time deployment capabilities (Ghorbani et al., 2020).

Upon assessing the LSTM-based approaches alongside different methods of performing sentiment analysis, several unique pros and cons begin to surface. LSTM models do not follow the same lexicon-sentiment based approaches, as those use an emotional dictionary; such models tend to learn emotion’s context from data itself, which also makes it easier to implement. This characteristic enables LSTMs to cope with varying patterns of language as well as new ways of expressing emotions (Yadav & Vishwakarma, 2021).

LSTMs outdo their peers that utilize basic Machine Learning algorithms like Naive Bayes or SVM in almost all aspects. They are especially more efficient with longer texts that are unstructured and compressed containing emotional hints distributed throughout the sequence. For instance, Zhu et al. (2022) proclaimed the high accuracy BiLSTM models achieved while identifying emotions from EEG signals illustrating the ability of these models to handle dependencies in a sequence.

Bi-Directional LSTMs (BiLSTM) improve performance even more because they analyze input sequences both forwards and backward; this captures a more advanced layer of nuanced emotional context. Liang et al. (2020) integrated a deep convolutional BiLSTM in facial recognition which increased the performance of the model as local feature extraction and contextual understanding captured expression at the same time.

In contrast to CNNs, LSTM-based models seem to perform better at capturing long-term dependencies in the data. On the other hand, CNNs perform better in computational speed and in spotting localized emotional features. Zhao et al. (2019) investigated a hybrid CNN-LSTM architecture for speech emotion recognition and proved that integration of both models results in more effective performance on both short-term and long-term features.

Despite this, the most significant limitation of LSTM-based models is their high computational burden. According to Masuda & Ikuko Eguchi Yairi (2023), although CNN-LSTM architectures are effective for tasks like emotion classification based on physiological signals, they demand high amounts of computation and extended training times, which could pose problems in real-time or resource-limited environments.

Other than that, emotion detection sentiment analysis relies on the application of various algorithms (Medhat et al., 2014). Moreover, Naive Bayes is a probabilistic classifier that considers the features being independent and thus rather simplifies the classification task, although it may be disappointing as in real life the features do not function independently (Imran et al., 2020). SVM on the other hand is a very powerful classifier which performs categorization of data by ascertaining the best decision boundary of the data which makes it fit for data in high dimensions such as text (Huang et al., 2022). BERT is essentially a transformer model depicted as state-of-the-art has also become popular due the allowance of capturing bidirectional context which aids the model to comprehend quite intricate sentences and their emotions (Devlin et al., 2018). One of the advantages of the mentioned algorithms, however, is their shortcomings, Naive Bayes performs poorly when there are dependent features, SVM is resource costly, and BERT does not go untrained (Devlin et al., 2018).

There have been many advancements in making sentiment analysis systems. One of the most common language processing toolkits for Languages is NLTK (Natural Language Toolkit), a popular library for the Python programming language that offers capabilities such as text processing, tokenization, and sentiment analysis (Ao, 2019). However, while NLTK can be effective for small projects or educational purposes, it is likely not the best fit for larger-scale industrial use (Mohan, 2016). For such cases, it is better to use spaCy, another and faster NLP library as it is more tailored for production environments. It allows working with pretrained models from Hugging Face library such as BERT, GPT and others what is in demand when developing modern systems for sentiment analysis (Devlin et al., 2018). As a result, these tools, integrated with large databases such as MySQL or MongoDB, make it possible to create prototypes of sentiment analysis systems. For example, aimed at storing the structured data, MySQL is probably the most suitable, although not so easy to use with social media messages that are not structured, for which MongoDB is much more suitable (MongoDB, 2024, Barahona & Sun, 2017).

Different approaches and tools for emotion detection sentiment analysis has its advantages and disadvantages. Lexicon-based approaches such as VADER are simple and fast, but they are not very effective in dealing with complex language or context. Machine Learning models like Naive bayes and SVM are more flexible and perform better, but they are data intensive as well as computationally expensive. Deep learning techniques, especially transformer-based models such as BERT, achieve better performance in terms of both accuracy and contextual information but at the cost of heavy training requirements and lower interpretability. Both NLTK and spaCy have their strengths and weaknesses in terms of text processing, with NLTK being easier to learn for starters while spaCy providing more advantages for powerful processing needs. Additionally, Hugging Face's Transformers library allows sentiment analysis systems to complete further enhancements as it allows access to a variety pre trained models although it is resource hungry.

### Summarization Comparison Table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No.** | **Approach/Tool** | **Pros** | **Cons** | **Best Use Case** |
| 1. | Lexicon-based Methods (example: VADER) | - It is simple and easy to implement.  - It has real-time processing capabilities.  - It is effective for informal text like social media posts. | - It is poor at handling context, sarcasm, and idioms.  - It has limited to predefined lexicons.  - It struggles with domain-specific vocabularies. | It can be used in social media monitoring and tasks requiring speed over accuracy. |
| 2. | Machine Learning Models (example: Naïve Bayes, SVM and Random Forest) | - It has more flexible and robust.  - It is suitable for labelled datasets.  - It is capable of detecting patterns in data.  - It has ensemble methods to improve accuracy. | - It is costly to compute.  - It requires labelled data.  - SVM and Random Forest are resource-intensive for large datasets. | It can be used in text classification tasks with structured datasets. |
| 3. | Deep Learning Models (example: RNN, LSTM, BERT) | - It has a high accuracy and ability to capture context.  - It is effective at handling complex language structures like idioms and sarcasm.  - It is suitable for large datasets.  - BERT is a pre-trained model which can reduce development time and cost. | - It requires a lot of resources, including GPUs and TPUs.  - It requires less interpretable.  - It needs a higher training and maintenance costs.  - The real-time analysis may occur delays without robust infrastructure. | It is suable to used for complex emotion detection tasks which require high accuracy and contextual understanding. |
| 4. | Rule-based methods (Example: Lexical Affinity, Keyword-based) | - It is easy to interpret.  - It is suitable for small datasets.  - It has low cost. | - It has limited scalability.  - It struggles with multi-label emotions and inherent complexities of text. | Limited domains with defined rules and lexicons. |
| 5. | Libraries/Tools (example: NLTK, spaCy, Hugging Face Transformers) | **i. NLTK**  - It is suitable for small projects.  **ii. spaCy**  - It produced faster and ready for production.  **iii. Hugging Face**  - It has the access to pretrained models which available for cutting-edge performance. | **i. NLTK**  - It is not scalable for industrial use.  **ii. spaCy and Hugging Face**  - It is complex for the beginners. | Text processing for sentiment analysis in both research and production environments. |
| 6. | Databases (example: MySQL, MongoDB) | **i. MySQL**  - It is good for structured data.  **ii. MongoDB**  - It handles unstructured text effectively like social media messages. | **i. MySQL**  - It is difficult for unstructured data.  **ii. MongoDB**  - It may lack efficiency for complex queries. | The storage is used for sentiment analysis which store structured data in MySQL while unstructured data in MongoDB. |

**Table 2.1 Comparison of different types of algorithms and tools**

Based on Table 2.1, it gives a basic outline of the various methods and tools utilized in sentiment and emotion recognition starting with yin and rule and ending with databases euro bank of ideas. The emphasis lies on their benefits, drawbacks and optimum application for lexicon approach, machine learning models, deep learning models, or rule-based techniques, or which library. The table is intended to streamline the most appropriate method based on the project’s requirements in terms of precision, flexibility, expense, and intricacy of implementation.

### Comparison Table of Algorithms

|  |  |  |  |
| --- | --- | --- | --- |
| **Algorithm** | **Lexicon-based Methods** | **Machine Learning Models** | **Transformer-based Models** |
| **Core Functionality** | Matching sentiment words according to rules. | Classification for text analysis based on statistics. | For text classification, contextual comprehension and fine turning are essential. |
| **Accuracy** | Low to moderate, depends on dictionary quality. | Moderate for simple sentiment classification tasks. | High.  It particularly for contextual and intricate tasks. |
| **Scalability** | Scalable.  Limited to rule complexity. | Restricted scalability.  It is slower when dealing with big datasets. | Highly scalable using distributed training. |
| **Multilingual Support** | Poor because each language requires a dictionary. | Limited because it requires separate models for different languages. | Excellent with pre-trained multilingual models. |
| **Resource Requirement** | Low. It is computationally inexpensive. | Moderate. It can run on standard CPUs. | High. It requires GPUs for fine-turning and inference. |
| **Example** | VADER, SentiWordNet | Naïve Bayes, SVM, Random Forest | BERT, RoBERTa |

**Table 2.2 Comparison of different types of algorithms**

In Table 2.2 are presented three algorithms that are most commonly used to perform sentiment analysis - lexicon-based methods, use of machine learning models and transformer-based models. It tells about their core functions, accuracy, scalability, support for multiple languages, resource requirements and illustrations. This comparison intends to assist in choosing the most suitable agent depending on the project requirements such as computing resources available, language range, and difficulty of the task to accomplish.

### Comparison of Tools

|  |  |  |  |
| --- | --- | --- | --- |
| **Tool** | **spaCy** | **NLTK** | **Hugging Face Transformers** |
| **Core Functionality** | Text preprocessing, tokenization and POS tagging. | Text preprocessing, tokenization and linguistic analysis. | Pre-trained transformer models for NLP tasks. |
| **Scalability** | Scales with preprocessing pipelines. | Scales with batch preprocessing. | Highly scalable with integration into ML workflows. |
| **Multilingual Support** | Strong.  - It supports multiple languages. | Moderate.  - It has less efficient for multiple languages. | Excellent.  - It supports multilingual models like BERT. |
| **Ease of Use** | Easy.  - It has a user-friendly API. | Moderate.  - It requires some scripting. | Moderate.  - It requires ML expertise. |
| **Resource Requirement** | Low | Low | High  - It needs GPUs. |

**Table 2.3 Comparison of different types of tools**

In Table 2.3, there are comparison of three editing as well as computer aided translation tools that are popular in the industry which are spaCy, NLTK and Hugging Face Transformers. The comparison looks at the key features, inter-organizational extension, support for multiple languages, usability, and requirement of resources. Through this table it shows the role of each of these tools for different project requirements, from simple text preprocessing tasks like spaCy and NLTK to more sophisticated efforts like employing ready transformer models as in Hugging Face. It assists in the selection of the most appropriate tool in terms of the resources, languages and the complexity of the project.

### Comparison of Cloud Platforms

|  |  |  |  |
| --- | --- | --- | --- |
| **Cloud Platform** | **Features** | **Use Case** | **Tools supported** |
| **Amazon Web Services (AWS)** | - Offers AWS Comprehend for ready-made sentiment analysis solutions. - Scalable computational resources for deploying ML models. - Real-time data processing. | - Consumer sentiment monitoring  - Brand monitoring  - Patient feedback evaluation | - Tensor Flow  - PyTorch  - AWS SageMaker |
| **Google Cloud Platform** | - Includes Google Natural Language API for sentiment analysis. - Pre-built AI and ML services. - BigQuery for large-scale data analysis. | - Big data sentiment analysis - Reputation analysis of cloud services - Real-time streaming sentiment analysis | - BigQuery - TensorFlow - Google AI Platform |
| **Microsoft Azure** | - Azure AI and Cognitive Services for sentiment analysis. - Seamless integration with Azure Machine Learning and distributed AI frameworks. | - Social media brand monitoring - Sentiment prediction of cloud services - Medical sentiment analysis | - Azure Machine Learning - PyTorch - ONNX |
| **Hybrid Cloud Models** | - Combines on-premises and cloud-based infrastructure. - Supports hybrid deep learning models like CBiLSTM for resource optimization. | - Automatic evaluation of other cloud services’ reputations. - Hybrid e-learning sentiment analysis systems. | - TensorFlow - PyTorch - Custom ML Frameworks |
| **Distributed ML Frameworks** | - Designed for large-scale data streaming and real-time analysis. - Cost-effective and highly scalable for big data applications. | - Streaming sentiment analysis - Scalable brand monitoring - E-learning feedback systems | - Apache Spark - TensorFlow - PyTorch |

**Table 2.4 Comparison of different types of cloud platforms**

This comparison looks at different cloud platforms that are designed for sentiment analysis of text and tools to detect emotions, covering their characteristics, tools they integrate with and areas of application. It details the functionalities of Amazon Web Services (AWS), Google Cloud Platform, Microsoft Azure, hybrid cloud models, and also distributed ML frameworks, illustrating their efficiency regarding large amount of data processing, real time data fetching, and Artificial Intelligence (AI) or Machine Learning (ML) technologies. The table also outlines the appropriate platforms and their features in handling tasks such as brand monitoring, patient feedback analysis and big data sentiment analysis.

## Feasibility Study

### Technical Feasibility

Technical feasibility assesses whether the available technology is sufficient to support the development of the emotion detection sentiment analysis system. The system requires robust Natural Language Processing (NLP) tools such as spaCy, NLTK, and Hugging Face Transformers, along with machine learning and deep learning frameworks like TensorFlow and PyTorch to process and analyze text data. Existing cloud infrastructure, such as Amazon Web Services (AWS) or Google Cloud, can provide the necessary computational resources for deploying deep learning models, particularly transformer-based models like BERT. Given the maturity of NLP technologies and the availability of scalable cloud solutions, this project is technically feasible. Additionally, databases like MySQL or MongoDB can efficiently manage large volumes of structured and unstructured data. However, the technical challenge lies in ensuring real-time analysis and multilingual support, which would require continuous improvements in data processing pipelines.

### Operational Feasibility

Operational feasibility examines the ability of the organization or team to operate and maintain the emotion detection sentiment analysis system. The system would need integration with real-time data sources, such as social media platforms, to monitor ongoing sentiment trends. Operational efficiency can be achieved by automating the processes of data collection, model retraining, and alert generation. However, the team must have expertise in machine learning, data science, and NLP to ensure smooth operations. Training end users, such as customer support teams or mental health professionals, to interpret the sentiment insights would also be crucial for operational success. Given the existing skills and frameworks, the system is likely operationally feasible, provided there is ongoing staff training and clear operational guidelines.

### Economic Feasibility

Economic feasibility refers to the cost-effectiveness of the project. The development of an emotion detection system requires initial investments in software, cloud computing, and data storage solutions, as well as ongoing costs for model retraining and maintenance. Open-source tools such as NLTK, spaCy, and pre-trained models from Hugging Face minimize initial software costs. Cloud providers offer scalable pricing models, allowing costs to grow alongside the project’s size. However, if real-time data processing and large-scale sentiment analysis are required, the costs could increase significantly. Economic feasibility is highly dependent on the project's scale and the intended volume of data processing. Overall, for smaller-scale operations, the system is economically feasible, but larger deployments may require careful cost management.

### Schedule Feasibility

Schedule feasibility assesses the time frame required to develop, test, and deploy the emotion detection sentiment analysis system. Development could be divided into phases: initial setup and data collection, model training, prototype testing, and deployment. The availability of pre-trained models and existing NLP frameworks significantly shortens the time required for development. For instance, leveraging LSTM for sentiment analysis allows faster prototype creation by skipping the need for training models from scratch. However, building custom datasets for specific emotional categories or languages may extend the timeline. A reasonable schedule for a prototype might range from 3 to 6 months, with additional time allocated for continuous improvement based on feedback and testing. Overall, the project is schedule-feasible if carefully planned with phased deliverables and resource allocation.

### Legal Feasibility

Legal feasibility involves analyzing the legal implications of developing and deploying the system. Since the emotion detection sentiment analysis platform will process potentially sensitive data from social media and online reviews, ensuring compliance with data privacy laws such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) is critical. The system must implement mechanisms to anonymize personal data and provide options for users to opt-out of data collection. Additionally, intellectual property considerations must be addressed, especially if third-party NLP tools and machine learning models are used. Licensing for cloud services and software tools must also comply with usage terms. Provided that these legal considerations are adequately addressed, the project is legally feasible.

## Chapter Summary and Evaluation

In recent years, the field of deep sentiment analysis and also emotion detection has gained significance along or combined with such areas as natural language processing (NLP) and machine learning which has its benefits in many spheres such as business, medicine and healthcare, education as well as social media analysis. For this purpose, the clients are given a lot of places to use. The chapter of this study aims at reviewing different approaches that include lexicon, machine learning, deep learning and many more sentiment analysis tools. VADER, SentiWordNet, and other lexicon methods are lighter and simpler than other methods for performing real-time sentiment analysis but laughable when it comes to detecting sarcasm, idioms, and other field-specific languages. For example, Naive Bayes, SVM or Random Forests models can gain high accuracy and reliability, but they require a lot of computing power as well as labeled data. In terms of model performance, deep learning models, recurrent neural networks (RNN), Long Short-Term Memories (LSTM), transformers (for instance BERT), have a strong hold on the performance metric such as language but have high hardware requirements. NLTK and spaCy, also Hugging Face Transformers library provide sentiment analysis with NLTK being used in educational projects while spaCy befitting for various industries. Although there is great development, the problems of emotions, informal language and specific cultural references remain unsolved. Following the paths of hybrid models as well as understanding context is likely to improve the performance of emotional race detection systems. This chapter shows well the advantages and disadvantages of different ways and stresses the importance of making further effort in closing the existing gaps and improving the performance parameters of the system.

The last part of the feasibility study investigates the potential for the implementation of the system for detection of emotions and sentiment across five dimensions. From a technical perspective, the project can be conducted due to the presence of modern NLP tools (for instance spaCy, NLTK, Hugging face transformers), machine learning frameworks such as TensorFlow and Pytorch as well as cloud offerings such as AWS or Google Cloud, though there are still some challenges such as real-time processing and multilingual support which require sound pipelines. From an operational perspective, the system can work with the proper integration of real time data, automation and machine learning and NLP teams with users who have been trained to understand the findings. From an economic perspective, using open-source applications and cloud applications reduces the start-up cost, making it possible for the smaller scale implementations, however, the large-scale operations require planning of the costs. Schedule feasibility is assured with pre-trained models and progressive development leading to prototype readiness between three and six months with extra period for enhancement. From a legal aspect of it, the project is plausible if data privacy laws such as the GDPR or CCPA are observed supported by data anonymization, opt-outs and proper licensing. In a nutshell the analysis substantiates the justification for the project, all other challenges are managed properly.

Chapter 3

Methodology and Requirements Analysis

# Methodology and Requirements Analysis

This chapter describes the methodology of the project and describes the algorithms as well as the system design that were developed during the research activities. It includes the software process model that was recommended, the functional and non-functional requirements of the system, and any other hardware or software requirements. The focus of the chapter is on the summary, analysis and assessment of the methodology.

## Overall Methodology

A diagram of a process

Description automatically generated

**Figure 3.1 Overall diagram of Methodology**

Based on Figure 3.1, there are various crucial stages in the process of creating the sentiment analysis system. During the Identification of Problem phase, a background study and literature review are done, and, in such cases, the research begins with a background study and literature review. This step involves defining the existing means of thinking about problems, algorithms, and tools for performing sentiment analysis such as lexicon approaches (for instance VADER), machine learning techniques such as Naive Bayes and SVM, and LSTM as deep learning models. The objective is to recognize deficiencies in available options and strategize the most suitable strategy for emotion detection to obtain desired results in terms of efficiency and context application. Moreover, the literature review also identifies the difficulties of determining sentiment in sarcastic or ambiguous statements, where conventional techniques might not be effective. This stage also emphasizes how crucial sequential memory capabilities which recurrent neural networks like LSTM offer for understanding emotional context in text.

The Planning phase is where the workflow of the project is structured. Some of the major activities in this phase are project scheduling which provides timelines for completion of deliverables and task assignment, where work is apportioned among team associates. This phase also encompasses the development of an appropriate resource activity plan to cater for management of all relevant activities from algorithm formulation to system implementation (deployment). Other than that, outlining the milestones for data collection, pre-processing, model training, fine-turning and evaluation is prepared. Certain risks are noted, for example, overfitting during model training or a lack of sufficient labelled data and strategies for their avoidance are formulated.

The first step in the Design and Development phase is requirements classification, which will reveal both functional and non-functional requirements of the system. The design of the system identifies key interfaces, like data input, data storage, data processing, information retrieval, and data output. The system’s design also focuses on scalability and flexibility for future upgrades and extensions. Text data is prepared for analysis employing data pre-processing approaches such as tokenization, stop words omission and lemmatization. Some of the special characters, URLs and excessive space are cleaned up as part of preprocessing. In addition to that, sourcing contractions which increases the extent of technical standardization is done for instance, “can’t” is expanded to “cannot”.

Due to the need to retain long-term dependencies and capture sequenced patterns in text, the LSTM model is used for sentiment analysis. It has to go through multiple procedures which makes it time-consuming to implement. To start, various data sources such as Facebook or Twitter as well as Yelp or Amazon reviews are collected along with datasets like Sentiment140, IMDB, or even Kaggle. After data accumulation, there comes the phase of data preprocessing and cleansing in which unwanted symbols such as “#”, emojis, and even the text case are removed. Changing the case of the text is done in order to make it uniform. The text is then split into sequences of words or subwords referred to as tokens, and these tokens are transformed into numbers through methods like word embeddings. Additionally, the sequences are padded to a fixed length which helps guarantee uniform input dimensions for the LSTM model and also enables processing of text inputs of varying lengths.

Apart from that, emotion detection tasks are done with Bidirectional LSTM models. First, embedding layer is made which is responsible for transforming words into dense vector representations, after that, Bidirectional LSTM layers process the text both from start to end and end to start so that both past and future tokens are captured. This includes adding dropout layers to avoid overfitting, as well as other methods like L2 regularization that are meant to improve generalization. It employs Adam optimizer as a means of optimization in the embedded system to mitigate the chances of over-fitting. To make sure that model does not overfit on the training set, it partitions the training data into two, one is a training set and the second is a validation set. With the use of TensorFlow and Keras libraries, the model is trained on local systems with GPU enabled or through cloud services such as Google Cloud. Alongside the optimizer, other parameters that include batch size, learning rate, and number of epochs are changed until the model is at a favorable level. When the results have reached a sustained value, that is when the validation phase the model undergoes is deemed to have begun. In the instance of a complete LSTM model, it combines into the environment supported by the Python scripting language, which is further interfaced with MongoDB database which eliminates possible conflicts within the data-model pre-processing stage, model prediction and output interpreted the construction phase.

Lastly, the evaluation stage allows for the validation of goals claimed by a given system. In order to calibrate the Forecasting model, real data sets are fed to it and its effectiveness is measured using different metrics like precision, recall, the F1 score, and execution speed. A confusion matrix is useful for showing how well the classification was done and the specific categories identified i.e. true positive, false positive, true negative and false negative. To ensure that the model is robust, test cases may have complex features like sarcasm, negations and idioms etc. To ensure accuracy, speed, and scalability, the engineers run the system through tests. All the functional and non-functional specifics are true and working so that the system is actually functional in the real world. After deployment the integration of model accuracy pipelines into real time operations together with the general retraining of the model to fit the changing linguistic shifts is done. This kind of approach makes it possible to follow certain sequence of actions while designing and deploying efficient sentiment analysis system that uses advanced technology and algorithms.

## Proposed Software Process Model

A diagram of a software development process

Description automatically generated

**Figure 3.2 Waterfall Model Diagram**

Based on Figure 3.2, the Waterfall model is found to be appropriate for this project since it is best suited for situations where the requirements are known from the start and few changes shall be made during project execution. The model is depicted as linear, and steps are performed sequentially because each step takes the output from the preceding one as input thereby developing in a systematic manner. It commences with the Requirement Analysis phase which deals with the gathering and documentation of the functional and non-functional requirements of the sentiment analysis system. For example, requirements for supporting emotion recognition, multi-language support, integration with hotel management platforms and high accuracy and scalability are finalized during this stage. These requirements act as the basis for all the remaining stages and once they have been documented, they are never updated.

The second stage is the System Design stage in which a more thorough description of the system’s architecture is developed. This is done through a breakdown of the major elements of the system such as text preprocessing, sentiment classification and visualization, and database storage modules. This phase also includes the selection of Python as the development language, TensorFlow and Keras for implementing LSTM models, and MongoDB for database storage. Deliverables include diagrams of the system architecture and thorough design specification that describe how the system is to be built.

Next, we will enter the implementation phase where the system is built in accordance with the specifications mentioned in the design documents. This includes the coding of the text preprocessing pipeline with such subtasks as tokenization and lemmatization, the embedding of the chosen LSTM-based sentiment analysis algorithm, and the creation of the results visualization user interface. The development is done in modular fashion so that the system can be easily maintained and expanded in future. After implementation, the system moves through Transition Testing that tests its usability and efficiency considering set requirements. It is validated through real world datasets to check its suitability, its responsiveness, and extensibility. Several metrics are used including precision, recall, and F1-score for the evaluation of the system’s performance so that both functional and non- functional criteria are achieved.

Lastly, the last step is the Deployment and Maintenance step where the system is released to the production environment. During maintenance some bugs that are reported because of normal use are fixed and there is a possibility of new updates or even changes being made due to user input or the environment. This phase helps to ensure the operational and effectiveness of the system for longer periods of time.

The Waterfall model is particularly appropriate for this project since it adopts a structured and concurrent approach. It is applicable in situations where requirements are stable, as is the case in most projects involving sentiment analysis where the task is to enhance emotion detection or embed a specific method. Another advantage of this model is its emphasis on documentation which allows the status of work, stakeholders, and the scope of all requirements to be controlled and covered comprehensively. The model possesses a strict structure that encompasses key deliverables including requirement specifications, design diagrams and test reports, which is perfect for quick completion of the project since the scope for subsequent changes is very low. On the other hand, had the requirements been anticipated to change or evolve, or the project integrated iterations, a more dynamic methodology such as Agile or Spiral would have been appropriate. For this project, because the scope is fixed and the objectives are set, the Waterfall model works perfectly.

## System Requirements

### Functional Requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **No.** | **Modules** | **Actors/Users** | **Requirements** |
| 1.0 | User Management | 1.1 End User | 1.1.1 The system should allow the user to register for an account in the sentiment analysis platform. |
| 1.1.2 The system should allow the user to log in and access their account securely. |
| 2.0 | Text Input and Preprocessing | 2.1 End User | 2.1.1 The system should allow the user to input text for sentiment analysis. |
| 2.1.2 The system should preprocess the text by removing stop words, tokenizing, stemming, and lemmatizing as required. |
| 3.0 | Sentiment and Emotion Detection | 3.1 End User | 3.1.1 The system should analyze the input text and identify the sentiment (e.g., positive, negative, neutral). |
| 3.1.2 The system should detect emotions (e.g., happiness, sadness, anger, fear) in the text. |
| 3.1.3 The system should display sentiment and emotion results with confidence scores. |

**Table 3.1 The list of functional requirements**

### Non-Functional Requirements

|  |  |  |
| --- | --- | --- |
| **No.** | **Category** | **Requirements** |
| 1.0 | Performance | 1.1 The system should process text inputs and return sentiment analysis results within 2 seconds. |
| 2.0 | Security | 2.1 The system should implement user authentication like multi-factor authentication to prevent unauthorized access. |
|  |  | 2.2 The system should implement input validation to prevent injection attacks and other security vulnerabilities. |
| 3.0 | Availability | 3.1 The system should achieve 99.99% uptime, ensuring uninterrupted access to users. |
| 4.0 | Usability | 4.1 The user interface should be intuitive and allow users to input text and view result without requiring prior training. |

**Table 3.2 The list of non-functional requirements**

## Other Requirements

### Hardware Requirements

|  |  |
| --- | --- |
| **Category** | **Requirements** |
| 1.1 Server | **- Processor:** Intel Xeon E5 or equivalent (minimum 8 cores, 16 threads) |
| **- Memory:** 32 GB RAM or higher |
| **- Storage:** 1 TB SSD for high-speed data access |
| **- GPU (for deep learning):** NVIDIA Tesla T4 or equivalent with CUDA support (for LSTM model fine-tuning) |
| 1.2 Client Devices | **- Processor:** Intel Core i5 or higher |
| **- Memory:** 8 GB RAM |
| **- Storage:** 500 GB HDD/SSD |
| 1.3 Networking | **- Internet Speed:** Minimum 100 Mbps for seamless client-server communication |
| **- Network Infrastructure:** Secure and reliable network with firewalls and VPN support |

**Table 3.3 The list of hardware requirements**

This table summarizes the hardware needs of the systems server’s client devices and networking infrastructure to ensure efficient operation. The server should be equipped with a good processing unit, such as the Intel Xeon E5 or its equivalent. It should also have not less than 32 GB RAM, 1 TB SSD storage to accommodate quick access to data and a GPU such as the NVIDIA Tesla T4 for tasks requiring deep learning such as fine-tuning of the LSTM model. Client devices need to be Intel Core i5 or octa core presentation enabled with 8 GB RAM and 500 GB of storage. Also, for effective and robust interaction between the client and server, at least 100 Mbps internet is necessary as well as networking with firewalls and VPN facilities for effective utilization.

### Software Requirements

|  |  |
| --- | --- |
| **Category** | **Requirements** |
| 2.1 Operating System | - Server: Ubuntu 20.04 LTS or Windows Server 2019 |
| - Client: Windows 10 or higher, macOS |
| 2.2 Programming Tools | **- Language:** Python 3.8 or higher |
| **- Libraries/Frameworks:** TensorFlow 2.x, Keras, NLTK, spaCy, scikit-learn |
| 2.3 Database | **- Type:** MongoDB (for unstructured data) or MySQL/PostgreSQL (for structured data) |
| **- Storage Requirement:** Minimum 500 GB database capacity |

**Table 3.4 The list of software requirements**

The table outlines the specific software requirements needed for the operation of the system, and operating systems, programming tools, and database specifications are presented in detail. In terms of the operating system, the server must use either Ubuntu 20.04 LTS or Windows Server 2019, while client machines may use Windows 10 and above or macOS. The programming tools involve utilizing Python 3.8 or higher and relevant libraries and frameworks which include but are not limited to: TensorFlow 2.x, Keras, NLTK, spaCy and scikit-learn. Database requirements suggest the usage of MongoDB to store unstructured data and MySQL or PostgreSQL for structured data within a relevant range of 500 GB or more to accommodate the data comfortably.

## Chapter Summary and Evaluation

This chapter explains the entire stages related to designing, implementing and maintenance of the Emotion Detection Text Sentiment Analysis System particularly in the case of following the Waterfall model which is preferred because of its orderly and linear nature throughout the process. The first part is structures from the literature review and the context in which different techniques of sentiment analysis such as VADER, Naive Bayes, SVM, LSTM, among others are described in terms of their potential for emotions detection in text. This analysis also points out some research deficits that exist and finds LSTM to be the most appropriate for the purpose of the system since it is good at comprehending sequential patterns and maintaining context over longer text sequences, making it suitable for emotion analysis.

The planning phase contributes directly to the project in terms of how it will be organized and resources allocated. At this stage project timeframes, obligations of parties involved, and risks concerning the project are set. Various components of the system are described in simple functional blocks like user management, text pre- and post-processing management, emotion recognition and mapping processes, and results visualization. To facilitate the construction of the system and systems in general, important development applications like Tensorflow, Keras, Pythin, and MongoDB were selected. Other than that, the requirements are pre-specified for viewing the features required for the system like secure user setup, text input, emotion recognition and showing the result. The Non-functional requirements include factors like speed of working, security, scalability and usability among others that are also taken into consideration to achieve the system’s operational targets. For instance, the claim is that the system operates in a high-speed mode where an input is provided, and a result is recalled within two seconds.

Moreover, the LSTM model has complicated operations such as training and running, so it was decided to use high performance hardware like Intel Xeon processors and NVIDIA GPUs for these purposes, which will contribute to system flexibility. Apart from this, MongoDB is employed in order to facilitate non-structured data storage. There are also measures that are included in the system that are aimed at preventing wrongful access to information and protection of data control.

The last step of the approach deals with the verification and the sustainment measures. The systems are tested using actual data sets and the systems are said to be working when certain metrics like precision, recall and F1 score are above a given threshold. From the evaluations, it can also be possible to tell that the system would be able to meet its goals and that it would perform successfully across various conditions. For this project, the Waterfall Model is selected because of its procedural orientation and its focus on paperwork, which ensures that each stage is developed in completeness and systematically. Following this approach guarantees robust documentation which provides visibility and traceability of the system through the development process. Furthermore, by adhering to this model, the project obtains the required degree of accuracy, levelling, and ease of use which leads to a well built and easily manageable sentiment analysis system.

This chapter includes the comments and is responsive, it puts stress on thorough planning, well-defined system architecture and novel instruments especially LSTM for sentiment analysis. The methodological approach is consistent with project objectives and coverage in that it seeks to ensure adequate architecture of the development process at its each stage. The performance and low tolerance to failure of the system are adequately addressed by the employment of ideal machine and program tools.

In this case, the waterfall model was a suitable selection because the project had a clear outline, thus making it easier for the project to follow phases of development. The stress placed on them makes it easier to understand and clear expectations of what should be in each phase, thus making the process more traceable and manageable. System’s functional requirements such as user identification, emotion tracing and system’s efficiency, as well as non-functional requirements are all inclusive.

Furthermore, this chapter shows a lot of effort to make sure that they deliver a good product by looking into certain challenges, for example data management and model tuning as well as testing the system with realistic data sets. The emphasis on metrics such as precision, recall and F1-score assures the reader about the critical need for accuracy especially in the systems developed for emotion detection. More so, the consideration of ease of use of the system means that it will be used by different people.

In a nutshell, the chapter provides one of the most detailed yet easy to follow approaches to the construction of an integrated and high-capacity sentiment analysis system. The careful choice of methods, instruments and strategies shows the willingness to ensure that the system works as designed and also as much as possible allows for further enhancement and expansion in the future.

Chapter 4

System Design

# System Design

The design of the system is showcased in this chapter along with models that include algorithms and processes explained using multimedia such as activity diagrams and pseudocode. It deals with capacities, models of interaction such as use cases, models of behavioral such as state and sequence diagrams, and other designs such as interface and architecture. This chapter ends with a conclusion and approval pertaining to the design.

## Algorithm/Process Design

### Activity Diagram: Text Input and Preprocessing

A diagram of a flowchart

Description automatically generated

**Figure 4.1 Activity Diagram of text input and preprocessing**

Based on Figure 4.1, the activity diagram describes the procedures for the text input and preprocessing on the sentiment analysis interface. The process starts with the user on the landing page and proceeds to login or registration. Once logged in, the user goes to the dashboard and clicks on the text input menu. The system prompts the user to enter the text input while checking to see that it is compliant with any set guidelines. If the input is acceptable, the system first alters the text by drop words, breaking it down into segments, converting words into their root form, or into their dictionary forms and generally sanitizing the text for analysis. The process ends with a success message and the text is ready for analysis of sentiment. Clear prompts are provided throughout the process to assist the user in completing the tasks without any problems.

### Activity Diagram: Sentiment and Emotion Detection

A diagram of a diagram

AI-generated content may be incorrect.

**Figure 4.2 Activity Diagram of sentiment and emotion detection**

The activity diagram for the Sentiment and Emotion Detection processing steps is shown in Figure 4.2. The workflow initiates with the user providing data to the system. This data is processed using a model which has already been trained and attempts to predict the sentiment and classifies the emotion as well. After that, a confidence score is derived indicating how trustable the prediction is. Upon completion of analysis, the results page is shown to the system. At this point, the user can select two action points for further engagement: viewing the detailed analysis report which summarizes the sentiments and emotions detected therein or, the user is able to download the report as a PDF file for offline storage and record keeping. No matter the decision of the user, the system saves the results automatically in the history section for retrieval. In the case of system processing error, the system will display the message immediately instructing the user to try again ensuring response error handling mechanisms are in place. The entire outline of the workflow is designed to make sure there is no ambiguity with the smooth user journey through the system where they can analyze, effortless interaction, and complete actions while providing failsafe measures to reinstate if there are pauses during the process.

## Database Design/Structural Model

### ERD

A diagram of a computer

AI-generated content may be incorrect.

**Figure 4.3 ERD of text sentiment analysis**

The ERD outlines a system of Sentiment Analysis which has four basic entities namely, User Account, TextInput, Sentiment Analysis and History. The users can attach text inputs to their accounts which are then analyzed. The system automatically creates an analysis record which shows the emotions detected including the confidence score for each. Besides that, the users can view their analysis result, and it will save to the history. This diagram illustrates the correct and effective design of the structure of the system so as to make it fully functional and data responsive.

### Class Diagram

A diagram of a computer

AI-generated content may be incorrect.

**Figure 4.4 Class Diagram of text sentiment analysis**

Based on Figure 4.4, this system is meant for detection of emotion and sentiment associated with the text. There is an ability to register, login and upload text as well using the UserAccount and TextInput components. The text goes through Emotion Recognition processes for text where confidence is measured, and results are provided. The History is when the user analyzed the text and it will automatically save it to history so when the user wants to trace back, they can go to history and check again their result. The system enables an end-to-end solution, starting from entering the text through the interface, analysis of the text for sentiment and emotion, and finally producing the report.

## Interaction Model

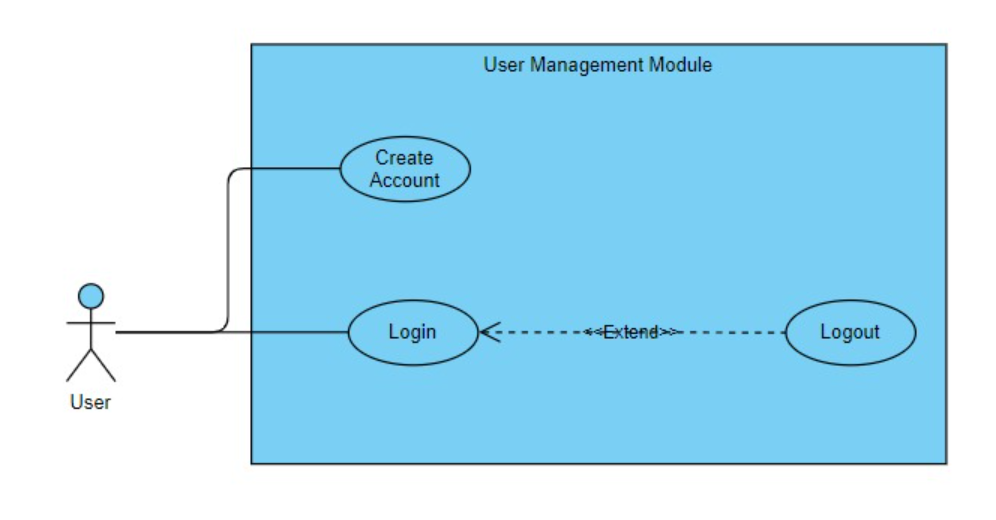
### Text Sentiment Analysis Subsystem Main Use Case Diagram

**A diagram of text and text

Description automatically generated**

**Figure 4.5 Main Use Case Diagram**

### Use Case 1 – User Management Module



**Figure 4.6 Use Case Diagram of User Management**

**Use Case Description Table**

**Table 1**

|  |  |
| --- | --- |
| **Use Case Name:** Create account | |
| **Actor:** User | |
| **Brief description:** This use case allows the user to register for a new account with the system. | |
| **Pre-condition:** - | |
| **Main Flow of Events:** | |
| **Actor Action** | **System Response** |
| 1. User navigates to the registration page. | 1. The system displays the registration form. |
| 1. User fills in the required information such as name, email, and password. 2. Select the “Submit” button to submit the registration form. | 1. The system validates the entered information, including the email format.   6) The account creation is successful if the details are valid and in the correct format.  7) The system directs the user to the login page. |
| **Alternative Flow of Events:**  A1. Step 4  If the user selects “Cancel” button, the registration form will be clear.  A2. Step 5  If the details are not valid and not in the correct format, the system will display an error message and ask the user to fill in again the details. | |

**Table 2**

|  |  |
| --- | --- |
| **Use Case Name:** Manage Account Information | |
| **Actor:** User | |
| **Brief description:** This use case enables the user to view and update their account information within the system. | |
| **Pre-condition:** The user must be logged in and have an active account. | |
| **Main Flow of Events:** | |
| **Actor Action** | **System Response** |
| 1. The user navigates to the account management section. | 1. The system displays the user’s profile information, including name, contact details and more. |
| 1. The user selects the option to edit the profile. | 1. The system allows the user to make changes to their profile information. |
| 1. The user saves the changes by confirming the update. | 1. The system updates the user’s profile information in the database. |
| 1. The user selects the option to change their password. | 1. The system prompts the user to enter their current password and then new password. |
| 1. The user confirms the password change. | 1. The system updates the user’s password in the database. |
| 1. The user selects the option to upload or update their profile picture. | 1. The system allows the user to upload or select a new profile picture. 2. The system updates the user’s profile picture in the database. |
| 1. The user logs out of their account. |  |
| **Alternative Flow of Events:**  A1. Step 5  If the user decides not to save the changes, the system discard the changes made by the user and returns to displaying the profile information.  A2. Step 9  If the user’s current password is incorrect, the system prompts the user to re-enter the correct current password before proceeding with the password change. | |

### Use Case 2 – Text Sentiment Analysis Module

A diagram of a text sentence analysis module

AI-generated content may be incorrect.

**Figure 4.7 Use Case Diagram of text sentiment analysis**

**Use Case Description Table**

**Table 1**

|  |  |
| --- | --- |
| **Use Case Name:** Analyze Text Sentiment and Emotion | |
| **Actor:** User | |
| **Brief description:** This use case allows the user input text and analyse its sentiment | |
| **Pre-condition:** The user must log in to the system. | |
| **Main Flow of Events:** | |
| **Actor Action** | **System Response** |
| 1. User navigates to the sentiment analysis homepage | 1. The system displays the dashboard with options to analyse text. |
| 1. User selects to upload text. 2. User clicks on the “Submit” button to process the uploaded text. | 1. The system validates the entered text input.   6) The system processes the text and classifies its sentiment as well as detect emotions.  7) The system displays the analysis result. |
| 8) User selects either to “View Report” or “Download Report” | 9) The system will provide the chosen output. |
| **Alternative Flow of Events:**  A1. Step 4  If the user selects “Cancel” button, the system stops the workflow and returns to the main screen.  A2. Step 5  If the uploaded text is in an invalid format, the system displays an error message and prompts the user to upload the correct format of text.  A3. Step 8  If the user does not select “View Report” or “Download Report”, the system keeps displaying the analysis result until the user decides. | |

## Behavioral Model

### User Management Module

UserAccount Class

A diagram of a login account

Description automatically generated

**Figure 4.8 State Diagram of User Management**

This state diagram represents the user account management. A user account is created in the Inactive Account state and when it is validated, it goes to Active Account state. From this point, the user has logged in and the system goes into the Logged In state. During this state, users can change their password by requesting a password change, moving to the Password Update state, and subsequently log in again when the password is successfully changed. It is also possible for a user to log out and then go to the Logged Out state which signifies the end of the process. This diagram describes the account lifecycle and its basic state transitions.

### Text Sentiment Analysis Module

TextInput Class

A diagram of a process flow

Description automatically generated

**Figure 4.9 State Diagram of Text Sentiment Analysis**

This state diagram depicts the activity of text sentiment analysis. It all starts with the receiving of input once the upload button has been clicked. Thus, the input is checked; if it is not ok, the system goes to the Rejected state, finalizing the process assuming it has been abandoned. Where the input is suitable, the minted text will go to Preprocessing, then the text will be subject to the Analyzing stage. Once the analysis is acceptable, the system proceeds to the Completed Analysis stage, while the results are waiting to be displayed in the Display Results stage. In the end, after presenting the results, the process of sentiment analysis is completed.

## Others Design - UI Design

A screenshot of a login page

AI-generated content may be incorrect.

**Figure 4.10 The main page of the sentiment analysis system**

This will be the main page of the text sentiment analysis system which can let the user to choose either login or create an account. If the user does not have an account, the user may click the “Create Account” button. Moreover, if the user has an account, the user can click “Login” button to access to the dashboard of the sentiment analysis system.

A screenshot of a login form

AI-generated content may be incorrect.

**Figure 4.11 The login page for user to enter username and password to login to the system**

This will be the login page which the user needs to key in their correct username and password to gain access to the sentiment analysis system. Once the user key in the correct username and password, it will automatically direct to the sentiment analysis page.

A screenshot of a login form

AI-generated content may be incorrect.

**Figure 4.12 The create account page for the new user to sign up**

For figure 4.12, this will be the create account page where user does not have an account yet. This is the page that user can fill in their details to create their new account. When the user key in wrongly like password or email, it will pop out error message and give an instruction to correct the format. After all the details are correct, the user may click “Register” button and the system will prompt a successfully message says that the user has done register and it will direct to the login page for user to login to the system. If the user presses wrongly, they can press the “Back to Login” button to return to the login page for login.

A screenshot of a computer

AI-generated content may be incorrect.

**Figure 4.13 The main dashboard for the sentiment analysis system**

This will be the main dashboard of the sentiment analysis system which can let the user to analyse their text to detect emotion and also for audio analysis as well.

A screenshot of a computer

AI-generated content may be incorrect.

**Figure 4.14 This will be the page that the user can input the text for preprocessing**

When the user press “Text Analysis” button in the main dashboard of the sentiment analysis system, it will proceed to this page. This can let the user to input the text for preprocessing and it will have an instruction button for user to know what the correct format is of inputting the text. Once the user inputted, the user may click the “Analyse Emotion” button and will proceed to the preprocessing page. Thus, if the user does not know what text they want to input, they can press “Try an example” button and the system will randomly input text for analyse emotion.

A screenshot of a computer

AI-generated content may be incorrect.

**Figure 4.15 This will be the processing page where the text is analysing**

This will be the preprocessing page where the page will display the details of processing and will direct to the sentiment analysis to detect emotion. Moreover, it also will calculate the confidence score for it.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.

**Figure 4.16 This will be the result page for the sentiment analysis**

After the preprocessing and sentiment analysis, it will display the result page for the user to view the result of the analysis and will have a description of what the emotion for the text that input by the user. Moreover, it will have highlighted the key emotional words and also visualization for the user to view.

A screenshot of a computer

AI-generated content may be incorrect.

**Figure 4.17 This will be the history page for the sentiment analysis**

After the user had done the analyse emotion, the user can click “History” button to view the page of the history. Moreover, the user can view the past emotion analysis, and it will just show up to last 10 analyses only.

A screenshot of a computer

AI-generated content may be incorrect.

A close-up of a document

AI-generated content may be incorrect.

**Figure 4.18 This will be the about page for the sentiment analysis**

The About Us page is where users can get to know what our sentiment analysis system is all about. This page is designed to give a clear and friendly explanation of what the system does, why we built it, and how it can be useful. Whether the user is curious or planning to use the platform, this page helps the user understand the goals behind our project, the technology powering it, and how it can make a difference.

## Chapter Summary and Evaluation

The Text Sentiment Analysis System developed focuses on the analysis of text inputs in order to analyse emotion while observing user and system requirements. The system is split into two major subsystems which are User Management and Text Sentiment Analysis with other features such as safe login, text input screening, preprocessing, and result output. In the User Management Module, the system requires secure credentials verification by confirmation of the user’s username and password. This module also includes error handling such as provision for number of failed attempts thus giving system-controlled access.

The Text Sentiment Analysis Module starts with the users providing their text input so that it can be validated and pre-processed. If the text fails basic validation, for example the text does not meet the acceptable formatting or content requirements, the process moves to an Error State or gets Rejected further. When the text gets everything right and successfully validated, it goes on to the Analysing phase where the system gets to do sentiment detection, meaning analysing the text provided in order to identify how the emotion is being expressed by the writer. When this is all done, the results indicating whether the sentiment was a positive, negative or no strong feeling of any type were displayed. This flow was effectively shown in the state flow diagram in which text input would ultimately transform into the display of the final analysis result.

Class diagrams were additionally created to represent the relationship between the important parts like UserAccount, TextInput, SentimentAnalysis and TextPreprocessing. The diagrams outline the relationships between the components: associations, aggregations and dependencies, so that the system structure and the data flow are clearly understood. Furthermore, simple and user-friendly graphics design was developed which consists of a login page for authenticating the user and a main page for text input and analysis results output. The usability of the UI in terms of error messaging and clarity is geared toward making use of the application efficient.

In evaluation, the system is fully functional because it provides a complete and secure user driven end-to-end sentiment analysis system. The transitions between the states allow the correct error and successful execution control flow while the class diagrams depict what is to be implemented structurally. The combination of graphic interfaces and detailed state diagrams contributes to the fact that the system is able to meet the necessary user needs and functional requirements. All in all, the project combines the usability of the system, protection of the system and accuracy of processed text sentiment to optimize the user experience.

Chapter 5

Implementation and Testing

# Implementation and Testing

This chapter details how the Emotion Detection Text Sentiment Analysis system was put into practice as well as the testing methods that were employed to confirm its functionality. It covers the essential elements of system implementation, such as model architecture, user interface development, and text preprocessing. A systematic test plan and comprehensive test cases are presented in the testing section in order to confirm the system's accuracy, usability, and performance in a range of scenarios.

## Implementation

The development of our Emotion Detection Text Sentiment Analysis system centers around three main pillars which are data preprocessing, model training, and delivering a user-friendly interface. Moreover, each of these components plays a crucial role in ensuring the system works efficiently and accurately. In this section, we’ll walk through the core features and highlight some of the most critical parts of the implementation, supported by relevant code examples to illustrate how everything comes together.

### Text Preprocessing

A computer screen shot of a computer code

AI-generated content may be incorrect.A computer screen shot of text

AI-generated content may be incorrect.

**Figure 5.1 Coding for text preprocessing**

In the preprocessing stage of sentiment analysis, it is critical for the data to be cleaned and formatted properly for the machine learning model to enable accurate functionality based on Figure 5.1. Such processes comprise removing unnecessary values like URLs or HTML tags, transforming special characters through character encoding, wiping out punctuation marks, and turning text into lowercase for uniformity before undergoing the expansion of contractions. Moreover, cleaning involves conversion into tokenized forms with the recognition that some non-essential words like “not” and “never” are vital for proper reasoning. Removing relevant words through lemmatization prompts words to be changed into base forms. Special negating words are particularly important to retain for sentiment analysis as they alter the meaning significantly. As such, the implementation utilizes a function-oriented manner focusing on numerous dimensions of text rationalization for accuracy. This serves the purpose of correcting discrepancies that result from the expansion of contractions.

### Bidirectional LSTM Model Architecture

A screenshot of a computer program

AI-generated content may be incorrect.

**Figure 5.2 Create Bidirectional LSTM Model**

Based on Figure 5.2, sentiment identification leverages a Bidirectional Long Short-Term Memory (BiLSTM) Neural Network Model for the sequential and contextual relationships in the text. The architecture has a word embedding layer at the beginning which transforms words into vectors, known as dense vectors, that contain semantic information; afterward, BiLSTM layers processes the text in a forward and backward manner to enhance context comprehension. The model uses dropout layers that switch off neurons with a certain probability for active neurons during training, and L2 regularization which penalizes larger weight and in this case both methods combat overfitting. The last dense layers are responsible for capture high-level features and conveying those features through softmax activation for classifying emotion into many categories. Other than that, considering structures from both sides tends to help much more when evaluating emotions than when considered alone since every word in the text has a surrounding context which helps determine the real meaning of the specific phrases.

### Emotion Analyze Class

A screenshot of a computer program

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

A screenshot of a computer code

AI-generated content may be incorrect.

**Figure 5.3 Emotion Analyze code**

The implementation contains an all-encompassing EmotionAnalyzer class which handles text analysis and presentation of results in a structured fashion which enhances ease of maintenance in Figure 5.3. This class incorporates various tools which work together: the BiLSTM deep learning model does the main emotion classification using a hierarchical approach to LSTM and VADER sentiment adds complementary rule-based sentiment scoring. This class also contain model integration for NLP with spaCy which processes linguistics data like named entities and dependencies which augments the understanding of the emotional content of the text. This approach encapsulates the entire emotion analysis workflow with relative ease and averting conflicting purposes through an object-oriented method which uses primary analysis alongside content deconstruction and graphical representation to blend into a single system that enhances usability while still enabling granularity.

### Web Application Implementation

A screenshot of a computer code

AI-generated content may be incorrect.

A screenshot of a computer program

AI-generated content may be incorrect.

**A white background with text

AI-generated content may be incorrect.**

A screenshot of a computer

AI-generated content may be incorrect.A screenshot of a computer

AI-generated content may be incorrect.

**Figure 5.4 Web Application Implementation**

According to Figure 5.4, the system is implemented as a Streamlit web application, which makes it easier for users to conduct emotion analysis through a simple point-and-click interface. The implementation has been done using a responsive UI that incorporates the columnar layout offered by Streamlit to ensure that user interfaces are pleasant and easy to interact with on various devices. In addition, robust state management enables the retrieval and storage of user data and selections through session variables. The application provides feedback through visual progress indicators so that users understand what the system is processing during analysis operations. Emotions detection and analysis are simplified through comprehensive result visualization where the system creates charts and graphs to portray emotions and their quantitative values which are difficult to comprehend. Moreover, the system features user registration with secure password storage and personalized history tracking so that users can monitor past analyses, enhancing the system’s effectiveness for continuous emotion assessment.

## Testing (Test Plan and Test Cases)

### Test Plan

### Overview

This test plan describes the approach for testing the Emotion Detection Text Sentiment Analysis System. It outlines scope, methodology, and criteria to be met in order to ensure the system’s functional and non-functional requirements are fulfilled. The plan encompasses manual and automated testing for the algorithms and user interfaces to verify system usability, performance, and detection accuracy of emotions.

### Context of Testing

Testing will take place under a controlled setting using the test datasets provided, user scenarios, and system configuration. Evaluation will focus on the functionality of the user interface and the machine learning model’s performance.

### Project

The Emotion Detection Text Sentiment Analysis System aims to analyze emotional content in text and audio data. Implementation involves deep learning models, especially classifying using labeled data through Biderectional LSTM neural networks. The emotions represented in the analysis include joy, sadness, anger, fear, surprise, disgust, shame, and other sentiments. The system offers a web-based interface developed with Streamlit, integrated with user authentication and skeleton history tracking functionalities.

### Test Item

|  |  |
| --- | --- |
| **Component** | **Description** |
| **Text Preprocessing** | Testing the feature that handles contractions, special characters and negation words in addition to cleaning and normalizing input text. |
| **Emotion Detection Model** | Confirming the precision and dependability of the BiLSTM model for textual emotion classification. |
| **User Interface** | Analyzing the functionality, responsiveness and user experience of Streamlit’s web application. |
| **User Authentication** | Validating the security and correctness of user registration, login and session management. |
| **History Tracking** | Testing the ability of users to store and retrieve prior analysis result. |
| **Fallback Analysis System** | Activation of the rule-based system when the machine learning model is not available. |
| **Data Visualization** | Emotion analysis results are tested using charts and visual aids. |
| **Error Handling** | Ensure that unintentional errors and invalid inputs are handled gracefully by the system. |

### Test Scope

**In Scope:**

1. **Functional Testing**

To guarantee comprehensive validation scope of the system, the testing will focus on multiple areas such as functional testing of the text analysis, which will check whether or not the system has the ability to perform text processing and classify the emotions accurately. In addition, robust testing of the system will involve providing samples of emotions encapsulated in elementary statements to complicated sentences which exhibit multi-layered emotions, therefore ensuring that the system performs well across various scenarios. The testing will also include accuracy testing where the detection model for emotion uses labeled datasets and is evaluated against an industry standard of at least 80% correct detections in different emotions.

1. **User Interface Testing**

UI testing will focus on the non-functional aspects of the web-app such as responsiveness and if it provides helpful information, as well as ensuring all text features are user friendly as users partake in navigating the system. This will include the checking of emotion analysis result visualization, validation of text input, and highlighting emotional words in the analyzed texts. Authentication testing is meant to validate the user registration and login processes with special concern to password encryption and other user protection layers against common attacks.

1. **Performance Testing**

The response time of the system with regard to text analysis will be tested to confirm that results are available within five seconds for normative inputs. Under Security Testing, it will be verified that the user information and text that has been analyzed is secure, particularly in regard to password security and session management.

**Out Scope:**

1. **Unit Testing**

Individual unit testing of the functions in the code base, stress testing with very large amounts of text, exploratory testing which is not bound by a specific test case or a defined set of pre-defined test cases, and regression testing of one or more previous versions are out of scope of this phase of testing. This also applies to automated tests done through CI/CD pipelines, penetration tests for evaluation of available security subsystems outside of basic authentication, and other non-standard tests not included in the current test plan.

### Assumption and Constraints

**Assumptions**

1. Assumed that emotion labeled test data is readily available for model validation.
2. The testing setup is expected to have the necessary GPU capabilities to support smooth and efficient model evaluation.
3. The system is designed specifically to analyze English text for emotion detection purposes.

**Constraints**

1. Some emotions may be underrepresented in the test dataset, which could affect the model’s ability to recognize them accurately.
2. Due to the tight project deadlines, there’s limited time available for thorough testing.
3. There are not enough resources to test the system’s performance across a wide range of platforms.

### Risk Register

**Product Risk**

1. **Model Accuracy Issues**

**Risk:** The text containing emotion may not be accurately classified within the emotion detection model.

**Mitigation:** Address misclassifications through identifying them using cross-validation and confusion matrix analysis.

1. **Data Privacy Issues**

**Risk:** Sensitive information may be embedded in user text inputs.

**Mitigation:** Ensure clear privacy information, limit history retention and implement secure storage.

**Project Risk**

1. **Technical Environment Issues**

**Risk:** Loading of the deep learning model may fail in production environment.

**Mitigation:** Set up automated error handling, fallback rule-based system and cover all boundaries designed.

1. **Timeline Pressure**

**Risk:** Equal partitioned testing if each emotion may take time.

**Mitigation:** Emphasis on testing coverage for common and bordering most emotions.

### Test Strategy

**Test sub-processes**

The testing will commence with a unit test of critical components which include the text preprocessing function and emotion detection model. Then the system will be integrated, and the interaction between the preprocessing module and the BiLSTM model along with the visualization module will be checked to confirm proper data flow. Then the system will be checked for the complete workflow validation of text sentiment analysis from input to result display and undergo user acceptance testing to check if the stakeholders’ expectations in terms of accuracy and usability are met.

**Test Deliverables**

Processes tested will be accompanied with a few key deliverables which document the process. All test cases outline specific contours which require scrutiny in text analysis with varying emotional input. In test execution, report passes and fails, record the ticks and time the tasks are marked completed. Test reports issue calculation errors in processing text and classifying emotion, performance measure response time to the text irrespective of the length of the text. A summary report incorporated all findings and provided recommendations.

**Test Design Techniques**

The testing will utilize specific black-box techniques geared to the challenges of emotion analysis. Equivalence partitioning will divide text samples into groups representing various emotions like joy, sadness, anger and more to ensure full coverage. Boundary value analysis will concentrate on texts with confidence scores bordering the thresholds that establish primary and secondary emotions. Decision table testing will check if emotion detection rules are correctly applied while negation words alter sentiment. Use case testing will confirm that the text analysis interface operates seamlessly through typical user interactions such as analyzing a text sample and viewing past analyses.

### Test Completion Criteria

1. All test cases related to text sentiment analysis have been executed.
2. High and critical severity bugs concerning the emotion detection feature have been fixed.
3. The detection of emotions continues to have at least an 80% accuracy threshold across all emotion categories.
4. Text analysis is performed reliably in less than 2 seconds.
5. Validation of user interface for each text input follows the expected workflows.
6. Every component responsible for rendering textual infographics accurately reflects relevant and correct information.
7. The secondary analysis model engages as intended when the primary model is offline.
8. All functionalities related to user accounts are implemented securely regarding signing up and signing in.

### Metrics to be collected

1. The ratio of executed test cases versus planned for text analysis features.
2. The density of identified defects per function/module.
3. Emotion detection accuracy metrics

* Precision, recall and F1-score for each emotion category
* Confusion matrix results for emotion classification
* Proportions for more than one concurrent sentiment classification

1. Performance metrics

* Average response time for text analysis
* Division of time allocated for preparing data, performing, processing and creating visualization output

1. User interface evaluation

* Proportionate measures of successful validation for interactive form fields.
* Overcoming errors

1. Test estimation information

* Evaluation of non-fulfillment of demands created on functionality of a product
* Quantitative measure of activity branches not simulated during enclosed user interactions

### Test Data Requirement

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Module** | **Function** | **Data** | **Data Type** | **Description** | **Sample** |
| Account | User  Sign Up | Username | string | Valid: The username contains the alphabet and number. | cookie11 |
| Valid: The username contains alphabet only. | cookie |
| Invalid: The username contains number only. | 123 |
| Invalid: The username contains special characters only. | ### |
| Invalid: The username field is empty. | Null |
| Password | string | Valid: The password contains the alphabet and number. | cookie123 |
| Valid: The password contains alphabet only. | cookie |
| Valid: The password contains number only. | 123456 |
| Invalid: The password is less than 6 characters. | 1234 |
| Invalid: The password field is empty. | Null |
| Confirm Password | string | Valid: The re-type password is the same as the password. | cookie123 |
| Invalid: The re-type password is not the same as the password. | cookie12 |
| Invalid: The confirm password field is empty. | Null |
| User Login | Username | string | Valid: The username matches the database. | cookie11 |
| Invalid: The username does not match the database. | cookie1 |
| Password | string | Valid: The password matches the database. | cookie123 |
| Invalid: The password does not match the database. | cookie12 |
| Text Sentiment Analysis | Text Input | Text | string | Valid: The text contains words and meaningful sentences. | "I feel great today!" |
| Valid: The text contains a mix of words and emoji. | "I'm so excited about my new job! It's a dream come true, and I couldn't be happier! 😄" |
| Valid: The text contains a mix of words and numbers. | "Today is 100% amazing!" |
| Invalid: The text field is empty. | Null |
| Invalid: The text contains only numbers. | "123456789" |
| Invalid: The text contains only special characters. | "@#$%^&\*" |

### Test Environment Requirement

**Hardware:**

1. Computing system with minimum 16GB RAM
2. GPU support for model testing
3. Standard desktop/laptop for UI testing

**Software:**

1. Python 3.8+
2. TensorFlow
3. Streamlit
4. Required Python libraries (NLTK, spaCy, etc.)
5. SQLite database
6. Jupyter Notebook
7. Visual Studio Code
8. Scikit-learn
9. Numpy
10. Pandas

### Testing activities and estimates

A grid of lines and squares

AI-generated content may be incorrect.

### Test Cases

**Overview**

To ensure that the Emotion Detection Text Sentiment Analysis System works as intended, we've put together a set of test cases that focus on its functionality, accuracy, and overall usability. Each test case is clearly identified and includes important details such as its purpose, any conditions that need to be met beforehand, the steps to follow during testing, what we expect to happen, and what happens when the test is run.

**Test Cases**

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_AccountModule\_UserSignUp\_001** | **Test Case Name: User Sign Up** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Account Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 20/3/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 22/3/2025** |
| **Short Description: Test the user sign up with a valid username, password, and confirm password.** | |

|  |
| --- |
| **Pre-conditions: The user does not have an account.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Click “Create Account” button | System navigates to the sign-up page. |  |  |
| 3 | Enter username: june11 |  |  |  |
| 4 | Enter password: june123 |  |  |  |
| 5 | Enter confirm password: june123 |  |  |  |
| 6 | Click “Register” button | System shows the message that the account was created successfully. |  |  |

|  |
| --- |
| **Post-conditions: The system navigates to the sign in page.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_AccountModule\_UserSignUp\_002** | **Test Case Name: User sign up** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Account Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 20/3/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 22/3/2025** |
| **Short Description: Test the user sign up with an invalid username, valid password, and confirmed password.** | |

|  |
| --- |
| **Pre-conditions: The user does not have an account.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Click “Create Account” button | System navigates to the sign-up page. |  |  |
| 3 | Enter username: |  |  |  |
| 4 | Enter password: june123 |  |  |  |
| 5 | Enter confirm password: june123 |  |  |  |
| 6 | Click “Register” button | System prompts an error message: “Please fill in all fields”. |  |  |

|  |
| --- |
| **Post-conditions: The system shows an error message, and the user is unable to sign up.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_AccountModule\_UserSignUp\_003** | **Test Case Name: User sign up** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Account Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 20/3/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 22/3/2025** |
| **Short Description: Test the user sign up with an invalid username, valid password, and confirmed password.** | |

|  |
| --- |
| **Pre-conditions: The user does not have an account.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Click “Create Account” button | System navigates to the sign-up page. |  |  |
| 3 | Enter username: 123 |  |  |  |
| 4 | Enter password: june123 |  |  |  |
| 5 | Enter confirm password: june123 |  |  |  |
| 6 | Click “Register” button | System prompts an error message: “Username must include letters”. |  |  |

|  |
| --- |
| **Post-conditions: The system shows an error message, and the user is unable to sign up.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_AccountModule\_UserSignUp\_004** | **Test Case Name: User sign up** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Account Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 20/3/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 22/3/2025** |
| **Short Description: Test the user sign up with an invalid username, valid password, and confirmed password.** | |

|  |
| --- |
| **Pre-conditions: The user does not have an account.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Click “Create Account” button | System navigates to the sign-up page. |  |  |
| 3 | Enter username: ### |  |  |  |
| 4 | Enter password: june123 |  |  |  |
| 5 | Enter confirm password: june123 |  |  |  |
| 6 | Click “Register” button | System prompts an error message: “Username must contain only letters & numbers”. |  |  |

|  |
| --- |
| **Post-conditions: The system shows an error message, and the user is unable to sign up.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_AccountModule\_UserSignUp\_005** | **Test Case Name: User sign up** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Account Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 20/3/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 22/3/2035** |
| **Short Description: Test the user sign up with valid username and confirmed password and invalid password.** | |

|  |
| --- |
| **Pre-conditions: The user does not have an account.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Click “Create Account” button | System navigates to the sign up page. |  |  |
| 3 | Enter username: june11 |  |  |  |
| 4 | Enter password: |  |  |  |
| 5 | Enter confirm password: june123 |  |  |  |
| 6 | Click “Register” button | System prompts an error message: “Please fill in all fields”. |  |  |

|  |
| --- |
| **Post-conditions: The system shows an error message, and the user is unable to sign up.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_AccountModule\_UserSignUp\_006** | **Test Case Name: User sign up** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Account Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 20/3/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 22/3/2025** |
| **Short Description: Test the user sign up with valid username, valid password and invalid confirmed password.** | |

|  |
| --- |
| **Pre-conditions: The user does not have an account.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Click “Create Account” button | System navigates to the sign-up page. |  |  |
| 3 | Enter username: june11 |  |  |  |
| 4 | Enter password: june123 |  |  |  |
| 5 | Enter confirm password: |  |  |  |
| 6 | Click “Register” button | System prompts an error message: “Password does not match”. |  |  |

|  |
| --- |
| **Post-conditions: The system displays an error message, and the user is unable to sign up.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_AccountModule\_UserSignUp\_007** | **Test Case Name: User sign up** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Account Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 20/3/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 22/3/2025** |
| **Short Description: Test the user sign up with a valid username, invalid password, and invalid confirm password.** | |

|  |
| --- |
| **Pre-conditions: The user does not have an account.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Click “Create Account” button | System navigates to the sign up page. |  |  |
| 3 | Enter username: june11 |  |  |  |
| 4 | Enter password: 12345 |  |  |  |
| 5 | Enter confirm password: 12345 |  |  |  |
| 6 | Click “Register” button | System prompts an error message: “Password must be at least 6 characters long”. |  |  |

|  |
| --- |
| **Post-conditions: The system displays an error message, and the user is unable to sign up.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_AccountModule\_UserLogin\_001** | **Test Case Name: User Login** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Account Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 20/3/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 22/3/2025** |
| **Short Description: Test the user login with a valid username and valid password.** | |

|  |
| --- |
| **Pre-conditions: The user must register as a user of the system.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Enter username: june11 |  |  |  |
| 3 | Enter password: june123 |  |  |  |
| 4 | Click “Sign In” button | System allows the user to log in. |  |  |

|  |
| --- |
| **Post-conditions: The system directs the user to the text sentiment analysis page.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_AccountModule\_UserLogin\_002** | **Test Case Name: User Login** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Account Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 20/3/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 22/3/2025** |
| **Short Description: Test the user with an invalid username and valid password.** | |

|  |
| --- |
| **Pre-conditions: The user must register as a user of the system.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Enter username: may |  |  |  |
| 3 | Enter password: june123 |  |  |  |
| 4 | Click “Sign In” button | System prompts an error message: “Invalid username or password” |  |  |

|  |
| --- |
| **Post-conditions: The system prompts an error message and the user is unable to log in.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_AccountModule\_UserLogin\_003** | **Test Case Name: User Login** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Account Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 20/3/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 22/3/2025** |
| **Short Description: Test the user with a valid username and invalid password.** | |

|  |
| --- |
| **Pre-conditions: The user must register as a user of the system.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Enter username: june11 |  |  |  |
| 3 | Enter password: june |  |  |  |
| 4 | Click “Sign In” button | System prompts an error message: “Invalid username or password” |  |  |

|  |
| --- |
| **Post-conditions: The system prompts an error message and the user is unable to log in.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_TextSentiment\_InputValidation\_001** | **Test Case Name: Text Input Validation** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Text Sentiment Analysis Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 22/3/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 23/3/2025** |
| **Short Description: Test the text input validation to ensure it accepts valid text and rejects invalid input** | |

|  |
| --- |
| **Pre-conditions: The user has accessed the emotion detection system and is on the text analysis input page.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Enter text: “I feel happy today !” | System receives the text. |  |  |
| 3 | Click “Analyze” button | System accepts the text and proceed to display the sentiment result. |  |  |

|  |
| --- |
| **Post-conditions: The system processes the valid input and returns the sentiment analysis result.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_TextSentiment\_InputValidation\_002** | **Test Case Name: Text Input Validation** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Text Sentiment Analysis Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 22/03/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 23/03/2025** |
| **Short Description: Test the text input validation to ensure it accepts valid text and rejects invalid input** | |

|  |
| --- |
| **Pre-conditions: The user has accessed the emotion detection system and is on the text analysis input page.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Enter text: "I'm so excited about my new job! It's a dream come true, and I couldn't be happier! 😄" | System receives the text |  |  |
| 3 | Click “Analyze” button | System accepts the text and proceed to display the sentiment result. |  |  |

|  |
| --- |
| **Post-conditions: The system processes the valid input and returns the sentiment analysis result.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_TextSentiment\_InputValidation\_003** | **Test Case Name: Text Input Validation** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Text Sentiment Analysis Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 23/03/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 24/03/2025** |
| **Short Description: Test the text input validation to ensure it accepts valid text and rejects invalid input** | |

|  |
| --- |
| **Pre-conditions: The user has accessed the emotion detection system and is on the text analysis input page.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | Enter text: "Today is 100% amazing!" | System receives the text |  |  |
| 3 | Click “Analyze” button | System accepts the text and proceed to display the sentiment result. |  |  |

|  |
| --- |
| **Post-conditions: The system processes the valid input and returns the sentiment analysis result.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_TextSentiment\_InputValidation\_004** | **Test Case Name: Text Input Validation** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Text Sentiment Analysis Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 23/03/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 24/03/2025** |
| **Short Description: Test the text input validation to ensure it accepts valid text and rejects invalid input** | |

|  |
| --- |
| **Pre-conditions: The user has accessed the emotion detection system and is on the text analysis input page.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | The user left empty field | System prompts an error message: “The text field cannot be empty” |  |  |

|  |
| --- |
| **Post-conditions: The system display an error message, and user cannot proceed to sentiment analysis until valid input is provided.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_TextSentiment\_InputValidation\_005** | **Test Case Name: Text Input Validation** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Text Sentiment Analysis Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 23/03/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 24/03/2025** |
| **Short Description: Test the text input validation to ensure it accepts valid text and rejects invalid input** | |

|  |
| --- |
| **Pre-conditions: The user has accessed the emotion detection system and is on the text analysis input page.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | The user enters numbers at the text field section | System receives the text |  |  |
| 3 | Click “Analyze” button | System prompts an error message “The text field cannot contain numbers only” |  |  |

|  |
| --- |
| **Post-conditions: The system display an error message, and user cannot proceed to sentiment analysis until valid input is provided.** |

|  |  |
| --- | --- |
| **Test Case Template** | |
| **Test Case #: TC\_TextSentiment\_InputValidation\_006** | **Test Case Name: Text Input Validation** |
| **System: Emotion Detection Sentiment Analysis** | **Module: Text Sentiment Analysis Module** |
| **Design By: Sandra Tang Poh Yi** | **Design Date: 23/03/2025** |
| **Executed By: Saw Hui Lin** | **Execution Date: 24/03/2025** |
| **Short Description: Test the text input validation to ensure it accepts valid text and rejects invalid input** | |

|  |
| --- |
| **Pre-conditions: The user has accessed the emotion detection system and is on the text analysis input page.** |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Step** | **Action** | **Expected System Response** | **Pass/Fail** | **Comments** |
| 1 | Open the emotion detection sentiment analysis system | System shows the sign in page. |  |  |
| 2 | The user enters only special characters at the text field section | System receives the text |  |  |
| 3 | Click “Analyze” button | System prompts an error message “The text field cannot contain special characters only” |  |  |

|  |
| --- |
| **Post-conditions: The system display an error message, and user cannot proceed to sentiment analysis until valid input is provided.** |

## Chapter Summary and Evaluation

The testing and implementation of the Emotion Detection Text Sentiment Analysis System showcase the development of a highly capable emotion analysis system. During implementation, various algorithms of natural language processing and deep learning were used to build an interactive interface which accurately detected emotions on text. In addition, the system boasts sophisticated preprocessing components and linguistics reasoning essential for emotion detection, a bidirectional LSTM based unit which permutes contextual information captured within relational texts, a strong rule-based backup system, advanced visualization systems, and the issuance of secured accounts with encrypted passwords. The evaluation procedure is organized in repeated incremental milestones with clearly defined end states; these include system functionality, performance benchmarking, and non-functional requirements like security and accuracy checking under various conditions and use cases. Future work could for example focus on improving understanding of sarcasm, cultural referencing, and adding support for real-time analysis. Overall, the system fulfills the goals of the project by enabling thorough automated detection of emotional values within blocks of text seamlessly through an intuitive interface and intelligently structured algorithms.

Chapter 6

Discussions and Conclusion

# Discussions and Conclusion

## Summary

The goal of the Emotion Detection Text Sentiment Analysis project is to automate the recognition of emotion in text content, a need which has only been growing. The system uses state-of-the-art machine learning Bi-directional Long Short-Term Memory (BiLSTM) neural networks to analyze and classify the text into different categories of emotions such as joy, sadness, anger, fear, surprise, disgust and shame. The project used elaborate text preprocessing methods to deal with contractions, special characters, and negation words that have great emotional context influence. The framework was implemented as a web application using Streamlit, allowing users to enter text, view analysis results, see visualizations, and maintain a history of previous analyses all in an effortless manner. There are graphs and visualizations which made the analysis much more interactive. The project was approached using Waterfall methodology which is a linear approach with distinct stages from requirement elicitation to deployment. This system represents a further step towards capturing the intricacies of human emotion entwined with text which can be used for mental health assessment systems and customer service systems, among other things.

## Achievements

The project has successfully accomplished its major milestones as stated in the proposal. To begin with, we researched emotions attached to the text and came up with an understanding of how different types of texts relate to different emotions. Secondly, we developed strong procedures for the collection, processing, and evaluation of text information. Our pipeline for text preprocessing effectively resolved problems related to contractions, special characters, and words of negation which are important for accurate emotion detection. Third, we designed and implemented a complete text-based sentiment analysis and emotion detection system with a friendly graphical user interface. The system goes beyond just primary emotions detection; it also detects other mixed emotions and provides confidence scores, thus deepening the emotional analysis. We also added advanced features that allow users to see and understand emotional data in simpler yet powerful ways. The system is more user friendly and secure with features like secure password encryption integrated to user authentication and history tracking functionalities. The results showed the system’s accuracy in emotion detection significantly enhanced as the target threshold of 80% was met for the different categories of emotions.

## Contributions

The Emotion Detection Text Sentiment Analysis system offers a powerful example of integrating technology into society. Technologically, The BiLSTM neural network for emotion detection is another application of deep learning in sentiment analysis. Its ability to identify mixed emotions and compute confidence levels adds complexity to the discipline does not present in simple sentiment classification into positive or negative. The hybridization of machine learning algorithms with rule-based frameworks creates a sturdy methodology that is useful in many areas. Socially, the system is helpful in monitoring mental health and, through text analysis, can usefully identify emotional distress at primary stages. This ability may assist in the early identification of so-called "crying for help" cases to guide people toward proper aid. In customer service, the system enables organizations to capture emotions from customers thereby using the correct response strategies, increasing customer satisfaction and loyalty. In education, technology provides instructors and teachers the ability to capture student's emotional responses to various instructional materials which may enhance educational outcomes. In addition, the system can display the results from emotion analysis in ways that do not require technical knowledge, thereby enabling more people to gain insights from emotion data analysis. Altogether, this project expands the existing literature in emotion detection which is the main goal of affective computing and enhances the available technology with practical application.

## Limitations and Future Improvements

The project has achieved a lot but it still has limitations which could be improved over time. One of the most notable limitations is the system’s language barrier, as it only performs text analysis in English at the moment. Future work on the system should aim at adding more languages, especially Malay and Chinese, so that the system becomes applicable in multi-lingual settings. The usage of sarcasm, cultural references that are laden with emotion, and context-sensitive expressions are also a problem for the system since more often than not, they get misinterpreted. Another concern the current model has is the use of idioms of figurative speech which may seem to express emotions that are not literally intended. Also, the emotions to be expressed are too many considering the context which makes the training data limited for the system.

In the future, adding more sophisticated models such as BERT would enhance understanding and accuracy. A more advanced mixed-emotion detection system would allow greater analysis of complex emotions and would allow greater analysis. Emotion detection in streaming data could also be implemented to make the system better for use in social media analytics. Building a more advanced analysis dashboard that allows users to visualize and analyze emotions over time would enable users to see changes in emotions as well. Moreover, adding an explainable AI system would allow users to understand how emotions were derived from the text. All of these changes, provided with great improvements, would enhance the system's accuracy and utility in emotion detection across various languages and contexts.

## Issue and Solutions

While developing the Emotion Detection Text Sentiment Analysis system, a range of challenges arose that required innovative approaches to solve them. One of the most prominent challenges was the accuracy of emotion analysis, especially when dealing with ambiguous and complex texts. The primitive model faced challenges in detecting some emotional subtleties such as in mixed emotions where contextual appreciation was key. To overcome this, we applied a more complex BiLSTM architecture and added more examples of emotional complexities to the training dataset. We also implemented a rule-based fallback approach together with the machine learning system which enhanced accuracy.

As with any other system, time was one of the challenges considering the system was being developed while the user was executing other academic assignments and other responsibilities. This meant that methodologies had to be concise and precise within defined timelines. In response to this, we employed modular development which allowed independent testing for components that could be developed concurrently. Project timelines were reprioritized to focus on essential elements with added enhancements later on when time allowed.

Throughout the development process, debugging constituted a never-ending problem, especially concerning the integration of various subsystems. Considerable effort was required to debug some issues, such as the memory leaks in the preprocessing pipeline or the inconsistencies in the visualization module. To address these problems, we adopted a more structured approach by developing a complete testing framework that included detailed test cases for every module, which in turn assisted with bug identification and resolution. Critical parts, such as the emotion detection model, were robust by using cross-validation strategies. Moreover, version control was implemented in order to manage changes and allow for rollback where needed, which was quite helpful during the debugging process. Regardless of all the issues, the challenges associated with resolving the problems in a systematic way eventually made possible a stable solution that fulfilled the goals of the project.

# References

‌Acheampong, F. A., Wenyu, C., & Nunoo‐Mensah, H. (2020). Text‐based emotion detection: Advances, challenges, and opportunities. *Engineering Reports*, *2*(7). <https://doi.org/10.1002/eng2.12189>

Aftab, S., Shah Muhammad, S., Awan, S., Ahmad, M., & Ahmad, S. (2017). Machine Learning Techniques for Sentiment Analysis: A Review. INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY SCIENCES AND ENGINEERING, 8(3). <https://www.researchgate.net/publication/317284281_Machine_Learning_Techniques_for_Sentiment_Analysis_A_Review>

Aggarwal, D., Bali, V., Agarwal, A., Poswal, K., Gupta, M., & Gupta, A. (2021). Sentiment Analysis of Tweets Using Supervised Machine Learning Techniques Based on Term Frequency. *Journal of Information Technology Management*, *13*(1), 119–141. <https://doi.org/10.22059/jitm.2021.80028>

Ahmad, Z., Jindal, R., Ekbal, A., & Bhattachharyya, P. (2020). Borrow from rich cousin: transfer learning for emotion detection using cross lingual embedding. *Expert Systems with Applications*, *139*, 112851. <https://doi.org/10.1016/j.eswa.2019.112851>

Ahmed, W. M. A. (2020). Stock market reactions to domestic sentiment: Panel CS-ARDL evidence. *Research in International Business and Finance*, *54*, 101240. <https://doi.org/10.1016/j.ribaf.2020.101240>

Al Ajrawi, S., Agrawal, A., Mangal, H., Putluri, K., Reid, B., Hanna, G., & Sarkar, M. (2021). Evaluating business Yelp’s star ratings using sentiment analysis. *Materials Today: Proceedings*. <https://doi.org/10.1016/j.matpr.2020.12.137>

Alaparthi, S., & Mishra, M. (2020). Bidirectional Encoder Representations from Transformers (BERT): A sentiment analysis odyssey. *ArXiv:2007.01127 [Cs, Stat]*. <https://arxiv.org/abs/2007.01127>

Alswaidan, N., & Menai, M. E. B. (2020). A survey of state-of-the-art approaches for emotion recognition in text. *Knowledge and Information Systems*, *62*(8), 2937–2987. <https://doi.org/10.1007/s10115-020-01449-0>

An algorithm and method for sentiment analysis using the text and emoticon. (2020). *ICT Express*, *6*(4), 357–360. <https://doi.org/10.1016/j.icte.2020.07.003>

Ao, S.-Iong. (2019). International MultiConference of Engineers and Computer Scientists : IMECS 2019 : 13-15 March, 2019 : The Royal Garden Hotel, Kowloon, Hong Kong. International Association of Engineers. <https://www.iaeng.org/publication/IMECS2019/IMECS2019_pp12-16.pdf>

Arcila Calderón, C., Ortega Mohedano, F., Álvarez, M., & Vicente Mariño, M. (2019). Distributed Supervised Sentiment Analysis of Tweets: Integrating Machine Learning and Streaming Analytics for Big Data Challenges in Communication and Audience Research. *Empiria. Revista de Metodología de Ciencias Sociales*, *42*, 113. <https://doi.org/10.5944/empiria.42.2019.23254>

Aslam, N., Rustam, F., Lee, E., Washington, P. B., & Ashraf, I. (2022). Sentiment Analysis and Emotion Detection on Cryptocurrency Related Tweets using Ensemble LSTM-GRU Model. *IEEE Access*, 1–1. <https://doi.org/10.1109/access.2022.3165621>

Barahona, J., & Sun, H.-M. (2017). A Process for Exploring Employees’ Relationships via Social Network and Sentiment Analysis. *Lecture Notes in Computer Science*, 3–8. <https://doi.org/10.1007/978-3-319-61845-6_1>

Batbaatar, E., Li, M., & Ryu, K. H. (2019). Semantic-Emotion Neural Network for Emotion Recognition From Text. *IEEE Access*, *7*, 111866–111878. <https://doi.org/10.1109/ACCESS.2019.2934529>

Bello, A., Ng, S.-C., & Leung, M.-F. (2023). A BERT Framework to Sentiment Analysis of Tweets. *Sensors*, *23*(1), 506. <https://doi.org/10.3390/s23010506>

Benedetto, F., & Tedeschi, A. (2016). Big Data Sentiment Analysis for Brand Monitoring in Social Media Streams by Cloud Computing. *Sentiment Analysis and Ontology Engineering*, 341–377. <https://doi.org/10.1007/978-3-319-30319-2_14>

‌Bhardwaj, A., Narayan, Y., Vanraj, Pawan, & Dutta, M. (2015). Sentiment Analysis for Indian Stock Market Prediction Using Sensex and Nifty. *Procedia Computer Science*, *70*, 85–91. https://doi.org/10.1016/j.procs.2015.10.043

Bharti, S. K., Varadhaganapathy, S., Gupta, R. K., Shukla, P. K., Bouye, M., Hingaa, S. K., & Mahmoud, A. (2022). Text-Based Emotion Recognition Using Deep Learning Approach. *Computational Intelligence and Neuroscience*, *2022*, 2645381. <https://doi.org/10.1155/2022/2645381>

Britzolakis, A., Kondylakis, H., & Papadakis, N. (2020). A Review on Lexicon-Based and Machine Learning Political Sentiment Analysis Using Tweets. *International Journal of Semantic Computing*, *14*(04), 517–563. <https://doi.org/10.1142/s1793351x20300010>

Castro, S., Hazarika, D., Pérez-Rosas, V., Zimmermann, R., Mihalcea, R., & Poria, S. (2019). Towards Multimodal Sarcasm Detection (An \_Obviously\_ Perfect Paper). *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. <https://doi.org/10.18653/v1/p19-1455>

Chiorrini, A., Diamantini, C., Mircoli, A., & Potena, D. (n.d.). *Emotion and sentiment analysis of tweets using BERT*. <http://star.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-2841/DARLI-AP_17.pdf>

Coviello, L., Sohn, Y., Kramer, A. D. I., Marlow, C., Franceschetti, M., Christakis, N. A., & Fowler, J. H. (2014). Detecting Emotional Contagion in Massive Social Networks. *PLoS ONE*, *9*(3), e90315. <https://doi.org/10.1371/journal.pone.0090315>

Cui, Q., Wu, S., Liu, Q., Zhong, W., & Wang, L. (2020). *MV-RNN: A Multi-View Recurrent Neural Network for Sequential Recommendation*. *32*(2), 317–331. <https://doi.org/10.1109/tkde.2018.2881260>

Davidson, G. P., Ravindran, D., Nirmal, V. J., & Pratheeba, R. A. (2024). Hierarchical Attention and Contextual Embedding for Robust Sarcasm Detection. *Indian Journal of Science and Technology*, *17*(40), 4243–4249. <https://doi.org/10.17485/ijst/v17i40.2682>

Desmet, B., & Hoste, V. (2013). Emotion detection in suicide notes. *Expert Systems with Applications*, *40*(16), 6351–6358. <https://doi.org/10.1016/j.eswa.2013.05.050>

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018, October 11). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. ArXiv.org. <https://arxiv.org/abs/1810.04805>

Freitag, M., Amiriparian, S., Pugachevskiy, S., Cummins, N., & Schuller, B. (2018). auDeep: Unsupervised Learning of Representations from Audio with Deep Recurrent Neural Networks. *Journal of Machine Learning Research*, *18*, 1–5. <https://www.jmlr.org/papers/volume18/17-406/17-406.pdf>

‌Garcia, K., & Berton, L. (2021). Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA. *Applied Soft Computing*, *101*, 107057. <https://doi.org/10.1016/j.asoc.2020.107057>

Garg, M., & Saxena, C. (2024). *Emotion detection from text data using machine learning for human behavior analysis* (pp. 129–144). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-443-22009-8.00008-2>

Ghorbani, M., Bahaghighat, M., Xin, Q., & Özen, F. (2020). ConvLSTMConv network: a deep learning approach for sentiment analysis in cloud computing. *Journal of Cloud Computing*, *9*(1). <https://doi.org/10.1186/s13677-020-00162-1>

Golder, S. A., & Macy, M. W. (2011). Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. *Science*, *333*(6051), 1878–1881. <https://doi.org/10.1126/science.1202775>

Huang, R., Wei, C., Wang, B., Yang, J., Xu, X., Wu, S., & Huang, S. (2022). Well performance prediction based on Long Short-Term Memory (LSTM) neural network. *Journal of Petroleum Science and Engineering*, *208*, 109686. <https://doi.org/10.1016/j.petrol.2021.109686>

Hung, L. P., & Alias, S. (2023). Beyond Sentiment Analysis: A Review of Recent Trends in Text Based Sentiment Analysis and Emotion Detection. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, *27*(1), 84–95. <https://doi.org/10.20965/jaciii.2023.p0084>

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1). <https://ojs.aaai.org/index.php/ICWSM/article/view/14550>

Imran, A. S., Daudpota, S. M., Kastrati, Z., & Batra, R. (2020). Cross-Cultural Polarity and Emotion Detection Using Sentiment Analysis and Deep Learning on COVID-19 Related Tweets. *IEEE Access*, *8*, 181074–181090. <https://doi.org/10.1109/access.2020.3027350>

Israel Edem Agbehadji, & Abosede Ijabadeniyi. (2021). *Approach to Sentiment Analysis and Business Communication on Social Media*. 169–193. <https://doi.org/10.1007/978-981-15-6695-0_9>

Jain, M. C., & Kulkarni, V. Y. (2014). TexEmo: Conveying Emotion from Text- The Study. *International Journal of Computer Applications*, *86*(4), 43–49. <https://www.ijcaonline.org/archives/volume86/number4/14977-3196/>

Jain, V. K., Kumar, S., & Fernandes, S. L. (2017). Extraction of emotions from multilingual text using intelligent text processing and computational linguistics. *Journal of Computational Science*, *21*, 316–326. <https://doi.org/10.1016/j.jocs.2017.01.010>

Kabir, M., Kabir, M. Md. J., Xu, S., & Badhon, B. (2019). An empirical research on sentiment analysis using machine learning approaches. *International Journal of Computers and Applications*, 1–9. <https://doi.org/10.1080/1206212x.2019.1643584>

Kramer, A. D. I. (2010). An unobtrusive behavioral model of “gross national happiness.” *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI ’10*. <https://doi.org/10.1145/1753326.1753369>

Krutilla, Z. (2023). Developing of NLP models by model based software developement. Gradus, 10(2). <https://doi.org/10.47833/2023.2.csc.024>

Li, S., Li, W., Cook, C., Zhu, C., & Gao, Y. (2018). *Independently Recurrent Neural Network (IndRNN): Building a Longer and Deeper RNN*. Openaccess.thecvf.com. <https://openaccess.thecvf.com/content_cvpr_2018/html/Li_Independently_Recurrent_Neural_CVPR_2018_paper.html>

Lopez Garcia, A., Tran, V., Alic, A. S., Caballer, M., Plasencia, I. C., Costantini, A., Dlugolinsky, S., Duma, D. C., Donvito, G., Gomes, J., Heredia Cacha, I., de Lucas, J. M., Ito, K., Kozlov, V. Y., Nguyen, G., Orviz Fernandez, P., Sustr, Z., Wolniewicz, P., Antonacci, M., … Plociennik, M. (2020). A cloud-based framework for machine learning workloads and applications. IEEE Access, 8, 18681–18692. <https://doi.org/10.1109/ACCESS.2020.2964386>

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment Analysis Algorithms and applications: a Survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. <https://doi.org/10.1016/j.asej.2014.04.011>

Min, W. N. S. W., & Zulkarnain, N. Z. (2020). Comparative Evaluation of Lexicons in Performing Sentiment Analysis. *Journal of Advanced Computing Technology and Application (JACTA)*, *2*(1), 1–8. <https://jacta.utem.edu.my/jacta/article/view/5207>

Mohaiminul, M., & Sultana, N. (2018). Comparative Study on Machine Learning Algorithms for Sentiment Classification. International Journal of Computer Applications, 182(21), 1–7. <https://doi.org/10.5120/ijca2018917961>

Mohan, V. (2016). *Text Mining: Open Source Tokenization Tools: An Analysis*. *3*(1), 37–47. <https://www.researchgate.net/publication/329800669_Text_Mining_Open_Source_Tokenization_Tools_An_Analysis>

MongoDB. (2024). *The most popular database for modern apps*. MongoDB. <https://www.mongodb.com/>

Najafabadi, M. K., Ko, C., Saman Shojae Chaeikar, & Shabani, N. (2024). A Multi-Level Embedding Framework for Decoding Sarcasm Using Context, Emotion, and Sentiment Feature. *Electronics*, *13*(22), 4429–4429. <https://doi.org/10.3390/electronics13224429>

Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, *11*(1). <https://doi.org/10.1007/s13278-021-00776-6>

Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, *11*(1). <https://doi.org/10.1007/s13278-021-00776-6>

Nasim, Z., Rajput, Q., & Haider, S. (2017). Sentiment analysis of student feedback using machine learning and lexicon based approaches. *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*. <https://doi.org/10.1109/icriis.2017.8002475>

Nguyen, H., Veluchamy, A., Diop, M., & Iqbal, R. (2019). Comparative Study of Sentiment Analysis with Product Reviews Using Machine Learning and Lexicon-Based Approaches. *SMU Data Science Review*, *1*(4). <https://scholar.smu.edu/datasciencereview/vol1/iss4/7/>

Opara, E. (2022). CLOUD-BASED MACHINE LEARNING AND SENTIMENT ANALYSIS. *Electronic Theses and Dissertations*. <https://digitalcommons.georgiasouthern.edu/etd/2515/>

Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., AL-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O., Hoste, V., Apidianaki, M., Tannier, X., Loukachevitch, N., Kotelnikov, E., Bel, N., Jiménez-Zafra, S. M., & Eryiğit, G. (2016, January 1). *SemEval-2016 Task 5: Aspect Based Sentiment Analysis*. HAL Archives Ouvertes. <https://doi.org/10.18653/v1/S16-1002>

Pranckevičius, T., & Marcinkevičius, V. (2017). Comparison of Naive Bayes, Random Forest, Decision Tree, Support Vector Machines, and Logistic Regression Classifiers for Text Reviews Classification. *Baltic Journal of Modern Computing*, *5*(2). <https://doi.org/10.22364/bjmc.2017.5.2.05>

Qaisi, L. M., & Aljarah, I. (2016). A twitter sentiment analysis for cloud providers: A case study of Azure vs. AWS. *2016 7th International Conference on Computer Science and Information Technology (CSIT)*. <https://doi.org/10.1109/csit.2016.7549473>

Rani, M., & Jan, S. (2020). Comparative Analysis and Enhancement of Sentiment Intensity Based Tools. *Open Source Systems*. <https://doi.org/10.1109/icosst51357.2020.9333119>

Rao, A., & Kishore Baglod. (2024). Role of sentiment analysis in education sector in the era of big data: a survey. *International Journal of Latest Trends in Engineering and Technology*, *Special Issue -*(SACAIM). <https://www.ijltet.org/journal_details.php?id=922&j_id=4018>

Razek, M. A., & Frasson, C. (2017). Text-Based Intelligent Learning Emotion System. *Journal of Intelligent Learning Systems and Applications*, *09*(01), 17–20. <https://doi.org/10.4236/jilsa.2017.91002>

Sailunaz, K., & Alhajj, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, *36*, 101003. <https://doi.org/10.1016/j.jocs.2019.05.009>

Sailunaz, K., & Alhajj, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, *36*, 101003. <https://doi.org/10.1016/j.jocs.2019.05.009>

Salam, S. A., & Gupta, R. (2018). Emotion Detection and Recognition from Text using Machine Learning. International Journal of Computer Sciences and Engineering, 6(6), 341–345. <https://doi.org/10.26438/ijcse/v6i6.341345>

Saleh, R. A., Driss, M., & Almomani, I. (2022). CBiLSTM: A Hybrid Deep Learning Model for Efficient Reputation Assessment of Cloud Services. *IEEE Access*, 1–1. <https://doi.org/10.1109/access.2022.3163246>

Salehinejad, H., Sankar, S., Barfett, J., Colak, E., & Valaee, S. (2018). *Recent Advances in Recurrent Neural Networks*. <https://arxiv.org/pdf/1801.01078>

Sangeetha, K., & Prabha, D. (2020). Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding for LSTM. *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-020-01791-9>

Sanyal, S., & Barai, M. K. (2021). Comparative Study on Lexicon-based sentiment analysers over Negative sentiment. *International Journal of Electrical, Electronics and Computers*, *6*(6), 01-13. <https://doi.org/10.22161/ijeec.66.1>

Singh, M., Jakhar, A. K., & Pandey, S. (2021). Sentiment analysis on the impact of coronavirus in social life using the BERT model. *Social Network Analysis and Mining*, *11*(1). <https://doi.org/10.1007/s13278-021-00737-z>

Singh, S., Kaur, H., Rubal Kanozia, & Kaur, G. (2023). Empirical Analysis of Supervised and Unsupervised Machine Learning Algorithms with Aspect-Based Sentiment Analysis. *Applied Computer Systems*, *28*(1), 125–136. <https://doi.org/10.2478/acss-2023-0012>

Sitendra Tamrakar, Bindu Madhavi, & Mohan, V. (2022). *Democratizing Sentiment Analysis of Twitter Data Using Google Cloud Platform and BigQuery*. 287–304. <https://doi.org/10.1002/9781119792642.ch14>

Tabinda Kokab, S., Asghar, S., & Naz, S. (2022). Transformer-based deep learning models for the sentiment analysis of social media data. *Array*, *14*, 100157. <https://doi.org/10.1016/j.array.2022.100157>

Torres, A. D., Yan, H., Aboutalebi, A. H., Das, A., Duan, L., & Rad, P. (2018). Patient facial emotion recognition and sentiment analysis using secure cloud with hardware acceleration. In Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications (pp. 61–89). Elsevier. <https://doi.org/10.1016/B978-0-12-813314-9.00003-7>

Wan, Y., & Gao, Q. (2015). An Ensemble Sentiment Classification System of Twitter Data for Airline Services Analysis. *2015 IEEE International Conference on Data Mining Workshop (ICDMW)*. <https://doi.org/10.1109/icdmw.2015.7>

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(55). <https://doi.org/10.1007/s10462-022-10144-1>

Wolny, W. (2016). Emotion Analysis of Twitter Data That Use Emoticons and Emoji Ideograms. *International Conference on Information Systems Development (ISD)*. <https://aisel.aisnet.org/isd2014/proceedings2016/CreativitySupport/5/>

Yue, T., & Jing, M. (2022). Sentiment Analysis Based on Bert and Transformer. *Springer Proceedings in Business and Economics*, 317–328. <https://doi.org/10.1007/978-3-030-75166-1_23>

Zad, S., Heidari, M., Jones, J. H. J., & Uzuner, O. (2021). Emotion Detection of Textual Data: An Interdisciplinary Survey. *2021 IEEE World AI IoT Congress (AIIoT)*. <https://doi.org/10.1109/aiiot52608.2021.9454192>

Zarra, T., Chiheb, R., Faizi, R., & El Afia, A. (2016). Cloud computing and sentiment analysis in E-learning systems. *2016 2nd International Conference on Cloud Computing Technologies and Applications (CloudTech)*. <https://doi.org/10.1109/cloudtech.2016.7847695>

Li, W., Qi, F., Tang, M., & Yu, Z. (2020). Bidirectional LSTM with self-attention mechanism and multi-channel features for sentiment classification. *Neurocomputing*, *387*, 63–77. <https://doi.org/10.1016/j.neucom.2020.01.006>

Kim, Y., Denton, C., Hoang, L., & Rush, A. M. (2017). Structured Attention Networks. *Arxiv.org*. <https://arxiv.org/abs/1702.00887>

Sailunaz, K., & Alhajj, R. (2019). Emotion and sentiment analysis from Twitter text. *Journal of Computational Science*, *36*, 101003. <https://doi.org/10.1016/j.jocs.2019.05.009>

Yadav, A., & Dinesh Kumar Vishwakarma. (2022). A Language-independent Network to Analyze the Impact of COVID-19 on the World via Sentiment Analysis. *ACM Transactions on Internet Technology*, *22*(1), 1–30. <https://doi.org/10.1145/3475867>

Zhu, M., Wang, Q., & Luo, J. (2022). Emotion Recognition Based on Dynamic Energy Features Using a Bi-LSTM Network. *Frontiers in Computational Neuroscience*, *15*. <https://doi.org/10.3389/fncom.2021.741086>

Liang, D., Liang, H., Yu, Z., & Zhang, Y. (2019). Deep convolutional BiLSTM fusion network for facial expression recognition. *The Visual Computer*, *36*(3), 499–508. <https://doi.org/10.1007/s00371-019-01636-3>

Zhao, J., Mao, X., & Chen, L. (2019). Speech emotion recognition using deep 1D & 2D CNN LSTM networks. *Biomedical Signal Processing and Control*, *47*, 312–323. <https://doi.org/10.1016/j.bspc.2018.08.035>

Masuda, N., & Ikuko Eguchi Yairi. (2023). Multi-Input CNN-LSTM deep learning model for fear level classification based on EEG and peripheral physiological signals. *Frontiers in Psychology*, *14*. <https://doi.org/10.3389/fpsyg.2023.1141801>

‌

‌